CERN-TH.4529/86

ANOMALIES, CHARACTERS, AND STRINGS

A.N. Schellekens

CERN, Geneva, Switzerland

N.P. Warner

Mathematics Department, Massachusetts Institute
of Technology, Cambridge, Mass., USA

A B S T R A C T

One-loop modular invariant heterotic and type II closed string theories are proved to be anomaly free, apart from terms proportional to $\text{Tr } F^2 - \text{Tr } R^2$. It is shown why this conclusion holds with remarkable generality for any conceivable fermionic string theory, regardless even of conformal invariance. A detailed discussion is given of the modular properties of character valued partition functions, upon which the proof is based. The fact that $\text{Tr } F^2 - \text{Tr } R^2$ terms remain in the fermionic contribution to the anomaly is shown to be a consequence of Quillen's holomorphic anomaly.

(Submitted to Nuclear Physics B)

CERN-TH.4529/86

October 1986

1. INTRODUCTION

It now seems clear that there is an intimate relationship between global anomalies on the string world sheet and local gauge and gravitational anomalies in the effective field theory of the string [1-3]. In this paper we will show that one-loop modular invariance of any 'heterotic' or 'type IIB' string in any (even) number of dimensions implies that the corresponding massless field theory has an anomaly that can be cancelled by the Green-Schwarz mechanism [4]. We find that in any dimension there are an infinite number of possible anomaly cancellable field theories. Moreover, the anomaly cancellation can be accomplished using the simplest possible Green-Schwarz mechanism, involving only a single, two-index antisymmetric tensor, $B_{\mu\nu}$. In other words, the anomaly from the fermion fields has a factor of $[\text{Tr}(F^2) - \text{Tr}(R^2)]$. It will also be shown that the fact that the anomaly from the fermion fields does not vanish entirely but factorizes in this manner, is due to the non-holomorphic factorization anomaly of Quillen [5].

In order to obtain the anomaly of the field-theory limit of a general heterotic string we will need to add together the contributions of fields in very different representations of the gauge and Lorentz groups. The first step in simplifying this problem is not to consider the complete Lorentz anomaly, but to consider the anomaly of the transverse, space-time rotation group. This neatly avoids the problem of having to deal with gauge fixing and ghost contributions to the anomaly. We postpone the justification for this method until Section 9.

Suppose that one has a fermion field ψ that transforms in a representation g of the gauge group and a representation $t \otimes S^{\pm}$ of the transverse rotation group, where t is a tensor representation and S^+ and S^- are the right- and left-handed spinors. (In this notation, taking t to be the vector representation would correspond to a Rarita-Schwinger field and a spinor field of opposite chirality -- the latter coming from the γ -trace.) The contribution of the field ψ to the gauge and gravitational anomaly is obtained from refs. [6-9].

$$A = \hat{A}(R) Ch(F, g) Ch(R, t) \quad (1.1)$$

where $\hat{A}(R)$ is the Dirac genus, R and F are the background gravitational curvature and Yang-Mills fields, and

$$Ch(F, g) = \text{Tr} \left[\exp \frac{iF_g}{2\pi} \right] \quad (1.2)$$

$$Ch(R, t) = \text{Tr} \left[\exp \frac{iR_t}{2\pi} \right]$$

are the Chern characters, in which F and R are to be viewed as 2-forms

$$F_g = \frac{1}{2} F_{\mu\nu}^a dx^\mu \wedge dx^\nu \Lambda_g^a$$

$$R_t = \frac{1}{2} R_{\mu\nu\alpha\beta} dx^\mu \wedge dx^\nu \Lambda_t^{\alpha\beta} \quad (1.3)$$

Here Λ_g^a is a representation matrix of the gauge group in the representation g , and $\Lambda_t^{\alpha\beta}$ is a transverse Lorentz rotation matrix in a representation t . In $(2N + 2)$ -dimensions, the contribution to the anomaly is obtained by applying the method of descent [8] to the $(2N + 4)$ -form in the expansion of eq. (1.1). From now on we will refer to this $(2N + 4)$ -form as the anomaly of the fermion field ψ .

Rather than calculating the complete anomaly for the massless level, it turns out to be simplest to calculate the anomaly generating function

$$A(g, F, R) = \hat{A}(R) \sum_{(g, t, m)} q^m Ch(F, g) Ch(R, t) \quad (1.4)$$

In a heterotic string a space-time fermion can only be obtained from the Ramond sector of the right-moving modes; the left movers only modify its gauge and Lorentz tensor structure. Indeed, a massless fermion only obtains gauge and further Lorentz

indices from the left-moving sector. Roughly speaking, one may associate the factor $\hat{A}(R)$ in eq. (1.4) with the right movers, and the other factor with the left movers. The sum is taken over all fields in representations (g, t) at each mass level m of the left-moving string. The anomaly is obtained by extracting the $(2N + 4)$ -form in the coefficient of q^0 . The function $A(q, F, R)/\hat{A}(R)$ is sometimes known as the character valued partition function of the string. In Section 2 we will make it plausible that in any sensible string theory of the heterotic type the anomalies are generated by a function of this form; in particular, we show that it cannot depend on $\bar{q}$, and that $A(q, 0, 0)$ should be a modular function of weight $-N$.

In Section 3 it will be shown how the Chern characters of fields in the gauge and transverse group representations arising in the string can be rewritten in terms of the Chern characters of the fundamental vector representations. This enables one to easily combine the contributions to the anomaly of all the different massless fermion fields.

These formulae are used in Section 4 to derive the anomaly generating function for any fermionic construction of heterotic strings, and to obtain the transformation properties of $A(q, F, R)$ under modular transformations.

Section 5 contains a brief discussion of the properties and classification of modular functions and modular forms. In Section 6, these results are applied to $A(q, F, R)$ to show that the anomaly factorizes.

Section 7 relates $A(q, F, R)$ to determinants of fermion operators on the string world sheet. In particular, we show that the modular property of $A(q, F, R)$ that results in the $[\text{Tr } F^2 - \text{Tr } R^2]$ factor in the anomaly arises from a logarithmically divergent Feynman diagram. It is the regularization of this diagram that is responsible for either Quillen's anomaly, or for the $[\text{Tr } F^2 - \text{Tr } R^2]$ factor in the field-theory anomaly.

In Section 8 we show how our results can be used to prove the absence of anomalies in type II strings. In Section 9 we motivate the use of light cone gauge in calculating gravitational anomalies. Although our main interest is, of course, in fully consistent string theories, we explain in Section 10 how many modular

invariant string theories, even otherwise inconsistent ones, may be used as a tool to construct anomaly-free massless field theories. We present the basic building blocks for constructing such theories, and give some examples. Section 11 contains some concluding remarks.

2. GENERAL DESCRIPTION OF HETEROTIC STRINGS

After the construction of the original $E_8 \times E_8$ or $SO(32)$ heterotic strings [10], several new ideas have emerged which generalize the original theory in various directions. These new ideas include mixing the modular transformations of left and right movers, as in the $O(16) \times O(16)$ string [11], various twistings of boundary conditions (orbifold compactifications [12] or changing the boundary conditions of some world sheet fermions [13], for example), Lorentzian self-dual lattices [14], and less healthy constructions yielding fermionic strings with tachyons, wrong statistics or lack of conformal invariance. In this section we will formulate the relevant properties of any meaningful generalization of the heterotic string in such a way that it becomes clear later why the results obtained in the rest of this paper (for the fermionic formulation of heterotic strings) can be expected to be valid in general.

The object of our study will be a closed string in an even number $(2N + 2)$ of dimensions. We will work in the light cone gauge formulation of the theory^{*)}, and assume that there are $2N$ left- and right-moving bosonic coordinates plus $2N$ right-moving fermionic ones. These are the only two-dimensional fields with space-time indices. There may also be a number of additional, but as yet unspecified, fields contributing to the various gauge symmetries. Although these assumptions may seem

*) Conformal invariance is needed to formulate the theory outside the light cone gauge. It is only in this sense that conformal invariance plays a role in our discussion.

to exclude type II strings, we will show in Section 9 that they can be included easily by turning some gauge symmetries into space-time symmetries [15].

The requirement of modular invariance for such a theory first arises at one loop, and is most conveniently discussed in Polyakov's path-integral formalism [16]. The one-loop diagram without external lines corresponds to an integral over all closed surfaces with the topology of a torus [17]. A fermion field on the torus can be periodic (+) or antiperiodic (-) along the two basic non-contractible loops on the torus. Altogether there are thus four possible boundary conditions [18-20], corresponding to the four spin structures on the torus. We denote them as (+ +), (+ -), (- -) and (- +) (with the first sign corresponding to the coordinate along the string), or alternatively as 1, 2, 3, 4, respectively (these numbers correspond to the four Jacobi θ -functions). To maintain transverse Lorentz invariance, the $2N$ space-time fermions should all have the same boundary conditions; the one-loop amplitude is then a sum over path integrals with these four fermion boundary conditions. In a theory of the type described above, the one-loop amplitude has the following form (see, for example, refs. [17-22])

$$\int \frac{d^2\tau}{(\text{Im } \tau)^2} (\text{Im } \tau)^{-N} P_B^N(q) P_B^N(\bar{q}) \sum_{i=1}^4 P_i^N(\bar{q}) T_i(q, \bar{q}) \quad (2.1)$$

with $q = e^{i\pi\tau}$, $\bar{q} = e^{-i\pi\bar{\tau}}$. The integration is over a complex parameter τ , with $\text{Im } \tau > 0$, which specifies the metric on the torus. The functions P_B^N and P_i^N are respectively, the partition functions for $2N$ bosons and $2N$ fermions with spin structure i , defined in Appendix A. For left movers they appear with argument q , for right movers with $\bar{q}$. The functions $T_i(q, \bar{q})$ are the partition functions of the part of the theory that we have not yet fixed.

The parameter τ on the full complex upper half plane gives a multiple cover of the space of all inequivalent metrics. This is because any modular transformation [21, 22]

$$\tau \rightarrow \frac{a\tau+b}{c\tau+d}, \quad a, b, c, d \in \mathbb{Z} \quad \text{with} \quad ad-bc=1$$

produces a metric which is equivalent to the one specified by τ . To get a meaningful answer, the integrand should be invariant under this transformation. In that case, one can restrict the τ -integration to a fundamental region, which covers the space of all inequivalent metrics precisely once.

The measure in (2.1) is already modular invariant by itself. Furthermore, we have the following transformations

$$\text{Im } \tau \rightarrow |c\tau+d|^{-2} \text{Im } \tau \quad (2.2)$$

the bosonic partition functions transform as follows:

$$\rho_B^N(q) \rightarrow \varepsilon(c\tau+d)^{-N} \rho_B^N(q) \quad (2.3)$$

where ε is a phase which cancels in (2.1). We see therefore that the bosonic part of (2.1) is also completely modular invariant by itself, so that the remaining factors must be fully modular invariant as well. This is true in any dimension. The factors $(\text{Im } \tau)^{-N}$ in (2.1) appear because N bosonic zero modes are factored out of the path integral (they provide an integral over transverse space time), and appear only for uncompactified bosonic coordinates. [For bosonic coordinates compactified on a lattice, $(\text{Im } \tau)^{-N}$ is replaced by a lattice partition function, transforming with the same weight.]

The three partition functions P_i with $i = 2, 3$, and 4 are transformed into each other by modular transformations, with certain phases, and hence the functions T_i ($i = 2, 3, 4$) must transform in the same way, with opposite phases. These sectors are of no interest to us, since they contain only non-chiral space-time fermions and bosons, and hence do not contribute to the anomaly. They correspond to the even

spin structures on the torus. The fourth partition function, P_1 , is obtained when all $2N$ fermions have periodic boundary conditions along both non-contractible loops on the torus. This corresponds to the odd spin structure. With these boundary conditions, the Dirac operator on the torus has a zero mode. Because the partition function is the determinant of this operator, it vanishes. Alternatively, one can understand this by observing that in this sector states with even and odd two-dimensional fermion number contribute with opposite signs. In particular the vacua, space-time Weyl spinors with opposite chirality, are related by an operator with two-dimensional fermion number 1, namely the zero-mode Ramond operator. Therefore, their contributions cancel. The same is true for all excitations of these vacua, so that the whole partition function vanishes. This structure of the vacuum is also precisely the reason why this sector is the only relevant one for a discussion of anomalies: in anomaly loops, two space-time fermions with opposite chirality, contributing with an *additional* opposite sign, are equivalent to two fermions with the same chirality, and the same sign.

Since the vacuum of the $(+ +)$ sector is the only source of chirality in the theory, the partition function multiplying it (or more precisely, the states contributing to that partition function) provides the gauge and Lorentz representations that may contribute to the anomaly. To determine these states we remove the multiplicative factor of zero due to the vacuum structure, or viewing $P_1(q)$ as a determinant, we factor out the fermion zero modes. This can be achieved by modifying the fermion boundary condition in the world-sheet time direction by a small phase $e^{i\nu}$, to remove the zero mode. The corresponding partition function is [20]

$$P_1^\nu(r, \bar{q}) = \left(\frac{\vartheta_1(r|\bar{q})}{\eta(\bar{q})} \right)^\nu$$

For small ν this can be written as

$$\rho'_N(\nu, \bar{q}) = \nu^N \left(\frac{g'_1(0|\bar{\tau})}{\eta(\bar{\tau})} \right)^N = (2\nu)^N \eta(\bar{\tau})^{2N}$$

The correct generalization of the modular transformations is the usual one for θ_1 (see Appendix A; note that the exponential factor is irrelevant for $\nu \rightarrow 0$). Therefore $\nu \rightarrow \nu/c\bar{\tau} + d$ in the transformation, so that $P_1^N(\nu, \bar{q})$ has weight zero. (We use the word 'weight' in a broader sense than usual, allowing for an overall phase in the modular transformations.)

We can now read off from (2.1) the function $P_A(q, \bar{q})$ that multiplies $(2\nu)^N$

$$P_A(q, \bar{q}) = P_B^N(q) P_B^N(\bar{q}) \eta(\bar{\tau})^{2N} T_1(q, \bar{q}) \quad (2.4)$$

From the modular invariance of (2.1), the transformation of $\text{Im } \tau$ [expression (2.2)] and the transformation of ν , we can conclude that P_A must be modular invariant with weight factor $(c\bar{\tau} + d)^{-N}$. The bosonic partition function is just the inverse of the η -function, so that the second and third factor cancel. There is a good reason for this cancellation. If this did not occur, the right-moving string would have massive excitations of the chiral ground state. This would produce chiral higher spin states for which there are no corresponding partners of opposite helicity. Such a space-time particle cannot be massive if Lorentz invariance is to be preserved. In general, one can expect this problem to arise if $P_A(q, \bar{q})$ depends on $\bar{q}$ ^{*}). Therefore $T_1(q, \bar{q})$, and hence $P_A(q, \bar{q})$, should be a function of q only.

^{*}) There is, in principle, a possibility that none of the massive chiral states satisfy the constraint $m_L = m_R$ between the left and the right sector. In other words, P_A could depend on $\bar{q}$ in such a way that there are no terms in which q and $\bar{q}$ have the same power. It is, however, practically inconceivable that there should be no way to excite the left or right sector to create massive chiral states with $m_L = m_R$. Notice, in particular, that the difference in the powers of q and $\bar{q}$ should always be even, to satisfy modular invariance for $\tau \rightarrow \tau + 1$, and that one can always bridge this difference by left-moving bosonic excitations.

This only says that at any level above the chiral ground state the multiplicities of the right-moving excitations cancel. It should be clear that a much stronger condition should be satisfied: there should not be any right-moving excitations at all in the chiral sector. In other words, for every excited state in a representation of the Lorentz group and additional gauge groups, there should be a state in the same representation, contributing to the partition function with opposite sign. In the known ten-dimensional strings this cancellation is due to world-sheet supersymmetry between periodic bosons and fermions, but in other cases, for example the models discussed in ref. [13], this is not so obvious.

The states of interest to us are thus the chiral vacuum of the right-moving string, multiplied with all states of the left-moving string^{*}). In this sector only the massless states are physical (i.e. satisfy $m_L = m_R$), but all others are needed for modular invariance. The partition function (2.4) contains only information about the multiplicities of these states at each level. In the following sections we will show how to modify it to get a more refined partition function, which provides the Chern characters of all gauge and Lorentz representations at each level. In some cases, P_A may vanish, but its character valued equivalent may be non-trivial. In those cases the properties of P_A derived above have only formal meaning, and anticipate similar properties of the character valued partition function.

Multiplication of the character valued partition function with the Dirac genus yields the anomaly generating function $A(q,F,R)$. As we will see, it inherits many of its properties from $P_A(q) = A(q,0,0)$. In particular, it should only depend on q , and transform with weight $-N$ (and an exponential factor) under modular transformations. The transformation of $A(q,0,0)$ is necessary, but not sufficient for that of $A(q,F,R)$. To derive the latter, one may either extend (2.1) to loop graphs with

^{*}) We will not consider the possibility that the right-moving ground state transforms in an additional gauge representation. Modular invariance does not appear to be of any use in controlling gauge and gravitational anomalies of this kind. We do not expect this situation to arise in a sensible string theory.

external lines or background fields [3], or, as we will do in this paper, construct $A(q,F,R)$ directly. In either case one has to define the string theory and its interactions more explicitly than we have done so far. In the next sections, we will consider theories constructed with left- and right-moving fermions with arbitrary boundary conditions (compactification of the left-moving bosonic string on self-dual lattices has been studied in ref. [2]). The transformation of $A(q,F,R)$ for these theories is given in Section 4. Because this transformation should also follow from the modular properties of loop graphs with external lines, we expect it to be valid for any sensible string theory, no matter how it is constructed.

3. CHERN CHARACTERS AND SYMMETRIZED TRACES

The Atiyah-Singer index theorem [6] gives the anomaly of any theory in terms of symmetrized traces of gauge and Lorentz representation matrices. In order to determine if there is an anomaly cancellation, one would like to express these traces in terms of the minimal number of basic invariants. A simple Lie algebra G of rank ℓ has ℓ such invariants. As a basis one may choose a set of traces of orders p_i ($i = 1, \dots, \ell$) over certain suitably chosen reference representations, or a set of ℓ symmetric tensors of rank p_i with gauge indices. The values of p_i are known for every G (see, for example, ref. [23]), and play an important role in the cohomology of Lie groups.

Defining $F_r = \Gamma_a \Lambda_r^a F^a$, where Λ_r^a is the set of representation matrices of G in the representation r , we can choose the following sets of basic traces for the classical Lie algebras [23]

$$\begin{array}{lll}
 A_n : & \text{Tr } F_V^k & k = 2, 3, 4, \dots, n, n+1 \\
 B_n, C_n : & \text{Tr } F_V^k & k = 2, 4, 6, \dots, 2n \quad (\text{even}) \\
 D_n : & \text{Tr } F_V^k & k = 2, 4, 6, \dots, 2n-2 \quad (\text{even}) \\
 & \text{and } \text{Tr } F_S^n
 \end{array}$$

Here 'V' denotes the vector representation and 'S' one of the two chiral spinor representations (the last trace changes sign when the chirality of S is flipped). The problem is to express traces of different order, as well as all traces over different representations in terms of sums of products of these basic traces. General expressions are not readily available, although special cases have been worked out with various methods [24, 25].

Chern characters are a useful tool for obtaining such expressions. The Chern character of the representation matrix F_r is defined as [26]

$$Ch(F_r) = \text{Tr} \exp\left(\frac{iF_r}{2\pi}\right) \quad (3.1)$$

The Chern character, being a sum of traces of powers of F_r , can of course be written as a sum of products of the basic traces. Among the terms in the sum one usually finds the ℓ basic traces themselves, with certain coefficients. These are the only terms that are not products of lower traces. Their coefficients are called the indices $I_k(r)$ [24]. For example, in $SU(n)$ one has

$$Ch(F_r) = \sum_{k=2}^n I_k(r) \frac{1}{k!} \text{Tr} \left(\frac{iF_r}{2\pi} \right)^k + \text{products of basic traces} + \dim(r)$$

Using the tensor product relation

$$Ch(\alpha \otimes \beta) = Ch(\alpha) Ch(\beta) \quad (3.2)$$

one can immediately derive a useful identity for the indices

$$I_i(r \otimes r') = \dim r' I_i(r) + \dim r I_i(r') \quad (3.3)$$

(This index sum rule, as well as an analogous one for subgroup embeddings, was first obtained with different methods in ref. [24].) In this section we will show how to

obtain in a systematic way all trace formulae for $SU(N)$, $SO(N)$, and $Sp(N)$.

Relations for exceptional algebras can often be obtained via their subalgebras. The problem consists of two parts; the first is to relate traces in arbitrary representations to the basic traces, defined above. This problem will be solved below. The second is to find expressions for very high order traces of the vector representation [for example, orders larger than N in $SU(N)$] in terms of lower order ones. These identities do not seem to play a role in string theory, and will be derived for completeness in Appendix B.

3.1 Trace formulae for tensor products

We denote a tensor power of a matrix M , symmetrized according to some Young tableau Y , as M_Y . Young tableaux are written either as $[k_1, \dots, k_a]$, where k_i is the length of the i^{th} column or $(\ell_1, \dots, \ell_b)$, where ℓ_i is the length of the i^{th} row. Using diagonalizable matrices one can easily prove the following identities [26]:

$$\det(1 + M) = \sum_{k=0}^{\infty} \text{Tr } M_{[k]} \quad (3.4)$$

$$[\det(1 - M)]^{-1} = \sum_{k=0}^{\infty} \text{Tr } M_{(k)} \quad (3.5)$$

where of course the first sum terminates at some value of k . We use these identities with $M = x \exp(iF/2\pi)$, where x is a parameter and $F = \sum_a F^a \Lambda_0^a$; Λ_0^a are the generators of a simple Lie algebra G in some representation r_0 . Again using diagonalizable matrices one can show

$$\text{Tr } (M_{[k]})^p = \text{Tr } (M^p)_{[k]} \quad (3.6)$$

and the analogous expressions for symmetric tensor products. Therefore, we obtain the following identities

$$\sum_{k=0}^{\infty} x^k \text{Ch}(F, [k]) = \det(1 + x e^{\frac{iF}{2\pi}}) \quad (3.7)$$

$$\sum_{k=0}^{\infty} x^k \text{Ch}(F, (k)) = \det(1 - x e^{\frac{iF}{2\pi}})^{-1} \quad (3.8)$$

Here $\text{Ch}(F, r)$ denotes the Chern character of F , evaluated in the representation r . By $\text{Ch}(F)$ we will mean the Chern character of the basic representation r_0 . These expressions are especially useful if r_0 is the vector representation of $SU(N)$. Then $[k]$ and (k) are irreducible representations, whose Chern characters can now be calculated by writing the determinants in eqs. (3.7) and (3.8) as the exponential of a trace of a logarithm, and expanding the logarithms.

$$\det(1 + x e^{\frac{iF}{2\pi}}) = \prod_{\ell=1}^{\infty} \exp \left[- \frac{(-x)^{\ell}}{\ell} \text{Ch}(\ell F) \right] \quad (3.9)$$

For example

$$\text{Ch}(F, [2]) = \frac{1}{2} [\text{Ch}(F)]^2 - \frac{1}{2} \text{Ch}(2F)$$

$$\text{Ch}(F, [3]) = \frac{1}{6} [\text{Ch}(F)]^3 - \frac{1}{2} \text{Ch}(2F) \text{Ch}(F) + \frac{1}{3} \text{Ch}(3F)$$

etc. Using the tensor product property of Chern characters one can obtain such expressions for any $SU(N)$ representation. For example, because of the tensor product relation

$$[k, \ell] = [k] \otimes [\ell] - [k+1] \otimes [\ell-1]$$

we can derive

$$\sum_{k,l} x^k y^l \text{Ch}(F, [k,l]) = \det(1 + x e^{\frac{iF}{2\pi}}) \det(1 + y e^{\frac{iF}{2\pi}}) (1 - \frac{x}{y})$$

(3.10)

To obtain the Chern character of $[k,l]$ in terms of the Chern character of $[1] = r_0$, one expands the right-hand side to the appropriate power in x and y , using eq. (3.9). Similar (though increasingly less practical) expressions can be derived for Young tableaux with an arbitrary number of columns. Thus the Chern character of any $SU(N)$ representation can be expressed in terms of polynomials in $\text{Ch}(\ell F)$, the Chern character of the vector representation.

From these expressions for $\text{Ch}(F, r)$ one can immediately read off the dimension of r [by setting $F = 0$ in the right-hand side, and using $\text{Ch}(0) = N$] and all indices, by selecting those terms on the r.h.s. that are not products of two or more traces. For example, the expression for $\text{Ch}(F, [2])$ given above can be written as

$$\begin{aligned} \text{Ch}(F, [2]) &= N \text{Ch}(F) - \frac{1}{2} \text{Ch}(2F) \\ &\quad + \text{products of basic traces} \\ &\quad + \frac{1}{2} N(N-1) , \end{aligned}$$

from which we read off, using the definition of the indices, that $I_k([2]) = N - 2^{k-1}$. As this example clearly shows, the only caveat is that $SU(N)$ has indices of order $2 \dots N$, so that traces of order higher than N should be expressible in terms of lower ones. The corresponding trace identities are derived in Appendix B.

All these expressions have an immediate generalization to row notation. In fact, it is easy to see that the Chern character polynomial for $(k_1 \dots k_n)$ can be derived from the one for $[k_1 \dots k_n]$ by replacing $\text{Ch}(\ell F)$ by $-\text{Ch}(\ell F)$, and adding an overall sign if the number of boxes of the Young tableau is odd.

3.2 SO(N) trace identities

Obviously any identity for tensor products of SU(N) vector representations is valid for tensor products of SO(N) vectors. These SO(N) tensor products may be reducible, but because the Chern character satisfies $\text{Ch}(\alpha \oplus \beta) = \text{Ch}(\alpha) + \text{Ch}(\beta)$ one can simply subtract traces from the Chern character. (This is never necessary in string theory.) Furthermore, some basic SU(N) traces may be decomposable in SO(N). Indeed, the basic traces are of order 2, 4, ..., N - 1 for N odd and 2, 4, ..., N - 2; N/2 for N even. If N is odd all odd traces are simply zero. The same is true if N = 4k, and in general for any tensor product of vector representations. The main problems are thus how to deal with spinor representations, with the trace of order N/2 (which exists only for spinor representations and self-dual tensors) and the absence of a fundamental trace of order N in SO(N), N even.

Consider first a fundamental spinor representation $S^\pm$ in SO(2M). In a basis where a (complexified) vector representation matrix $F_V/2\pi$ has eigenvalues $\pm iy_\beta$ ($\beta = 1 \dots M$), a Weyl spinor matrix $F_S/2\pi$ has eigenvalues $i/2 (\pm y_1, \dots, \pm y_M)$ with an even or odd number of - signs, depending on the chirality. Therefore one finds

$$\text{Tr} e^{\frac{iF_S}{2\pi}} = 2^{M-1} \left[\prod_{\beta=1}^M \cosh \frac{y_\beta}{2} \pm \prod_{\beta=1}^M \sinh \frac{y_\beta}{2} \right] \quad (3.11)$$

The first term can be expressed completely in terms of traces over the vector representation

$$\prod_{\beta=1}^M \cosh \frac{y_\beta}{2} = \exp \left[\sum_{n=1}^{\infty} \frac{(2^{2n}-1) B_{2n}}{4n (2n)!} \text{Tr} \left(\frac{iF_V}{2\pi} \right)^{2n} \right] \quad (3.12)$$

This formula can be derived easily from the Taylor series expansion of $\log \cosh x$; B_{2n} are the Bernoulli numbers. The second term can be written as follows:

$$2^{M-1} \prod_{\beta=1}^M \sinh \frac{y_{\beta}}{2} = \frac{1}{2} \prod_{\beta=1}^M y_{\beta} \exp \left[\sum_{n=1}^{\infty} \frac{B_{2n}}{4n} \frac{1}{(2n)!} \text{Tr} \left(\frac{iF_V}{2\pi} \right)^{2n} \right] \quad (3.13)$$

The prefactor cannot be written in terms of vector traces. The best one can do is call it a new fundamental trace. This term appears in eq. (3.11) with opposite sign for opposite chiralities. Therefore,

$$\prod_{\beta=1}^M y_{\beta} = \frac{1}{M!} \text{Tr} \left(\frac{iF_S}{2\pi} \right)^M \gamma_* \equiv \{\tilde{M}\} \quad (3.14)$$

(the projection matrices on the two chiralities are $1/2 (1 \pm \gamma_*)$ and $S = S^+ + S^-$).

Thus we find

$$\begin{aligned} \text{Ch}(F, S^{\pm}) &= \dim S^{\pm} \exp \left[\sum_{n=1}^{\infty} \frac{(2^{2n}-1) B_{2n}}{4n (2n)!} \text{Tr} \left(\frac{iF}{2\pi} \right)^{2n} \right] \\ &\quad \pm \{\tilde{M}\} \exp \left[\sum_{n=1}^{\infty} \frac{B_{2n}}{4n} \frac{1}{(2n)!} \text{Tr} \left(\frac{iF}{2\pi} \right)^{2n} \right] \end{aligned} \quad (3.15)$$

with all traces on the right-hand side in the vector representation. The same formula, but without the second term, holds for odd N . All higher spinor representations can now be obtained by means of tensor products. The trace relation of order $2M$ for $SO(2M)$ will be derived in Appendix B.

In index theorems one uses these Chern characters with gauge or Lorentz 2-forms as arguments, but apart from this difference all identities are exactly the same. Using the same methods one can derive

$$\hat{A}(R) = \prod_{\alpha=1}^M \frac{x_{\alpha/2}}{\sinh x_{\alpha/2}} = \exp \left[- \sum_{k=1}^{\infty} \frac{B_{2k}}{4k} \frac{1}{(2k)!} \text{Tr} \left(\frac{iR}{2\pi} \right)^{2k} \right] \quad (3.16)$$

where x_{α} are the skew eigenvalues of the curvature 2-form $R/2\pi$ in the vector representation of $SO(2M)$. For completeness we give also the formula for the Hirzebruch signature polynomial appearing in the self-dual tensor anomaly [8]

$$L(R) = \prod_{\alpha=1}^M \frac{x_{\alpha}}{\tanh x_{\alpha}} = \exp \left[\sum_{k=1}^{\infty} \frac{2^{2k} (2^{2k-1} - 1) B_{2k}}{2k (2k)!} \text{Tr} \left(\frac{iR}{2\pi} \right)^{2k} \right] \quad (3.17)$$

4. The anomaly generating functional

In this section we will consider a more concrete realization of the kind of theories considered in Section 2, namely models where gauge symmetries are generated by world-sheet fermions. For simplicity, we will first consider only periodic or antiperiodic boundary conditions for these world-sheet fermions. In the spirit of orbifold compactifications, one might consider more general 'fractional' boundary conditions as well, as was done in ref. [13]. We will show, at the end of this section, that the appropriate generalizations are straightforward.

As we argued in Section 2, the multiplicities of the states contributing to the anomaly at each level of the left-moving string are given by a modular invariant partition function $P_B^N(q) T_1(q)$. In a fermionic model this function is some linear combination of products of θ -functions, each term corresponding to a path integral with some choice of boundary conditions for each of the fermions. (Actually a θ -function corresponds to a pair of fermions with the same boundary conditions. In the following we will only consider pairs of fermions.) If k pairs of fermions have the same boundary conditions in every contribution to the path integral, the theory has an $SO(2k)$ global symmetry of rotations of these fermions into each other.

Depending on the precise construction of the theory, this symmetry may or may not be a local symmetry of the string. But, in any case, we can imagine coupling background fields to it, and investigate whether the symmetry is anomalous. To do that, we want to generalize the partition function to a character valued partition function. It is sufficient to consider groups of $2k$ fermions, all having the same boundary condition. The full partition function is then a sum of products of such combinations.

The (ordinary) partition functions of such groups can be written as follows:

$$\begin{aligned} \mathcal{P}_{--}^k(q) &= \text{Tr} [\exp(i\tau H_{NS})] \\ &= q^{-\frac{k}{12}} \prod_{n=1}^{\infty} (1+q^{2n-1})^{2k} = \left[\frac{\vartheta_2(0|\tau)}{\eta(\tau)} \right]^k \end{aligned} \quad (4.1)$$

$$\begin{aligned} \mathcal{P}_{-+}^k(q) &= \text{Tr} [(-1)^f \exp(i\tau H_{NS})] \\ &= q^{-\frac{k}{12}} \prod_{n=1}^{\infty} (1-q^{2n-1})^{2k} = \left[\frac{\vartheta_4(0|\tau)}{\eta(\tau)} \right]^k \end{aligned} \quad (4.2)$$

$$\begin{aligned} \mathcal{P}_{+-}^k(q) &= \text{Tr} [\exp(i\tau H_R)] \\ &= 2^k q^{\frac{k}{6}} \prod_{n=1}^{\infty} (1+q^{2n})^{2k} = \left[\frac{\vartheta_2(0|\tau)}{\eta(\tau)} \right]^k \end{aligned} \quad (4.3)$$

$$\mathcal{P}_{++}^k(q) = \text{Tr} [(-1)^f \exp i\tau H_R] = 0 \quad (4.4)$$

where H_{NS} and H_R are the Neveu-Schwarz and Ramond Hamiltonians, respectively, $q = \exp(i\pi\tau)$ and f is the world-sheet fermion number. The functions ϑ_i and η are defined in Appendix A. The partition function for $2N$ transverse bosons is

$$\mathcal{P}_B^N(q) = q^{-\frac{N}{6}} \prod_{n=1}^{\infty} (1-q^{2n})^{-2N} = \eta(q^2)^{-2N} \quad (4.5)$$

Using eqs. (3.7), (3.8), and (3.11), we find that the generating functions of the Chern characters of the gauge representations generated by the world-sheet fermions are simply the character valued partition functions analogous to eqs. (4.1)-(4.4); that is

$$P_{--}^k(q, F) = q^{-\frac{k}{12}} \prod_{n=1}^{\infty} \det(I + q^{2n-1} e^{\frac{iF}{2\pi}}) \quad (4.6)$$

$$P_{-+}^k(q, F) = q^{-\frac{k}{12}} \prod_{n=1}^{\infty} \det(I - q^{2n-1} e^{\frac{iF}{2\pi}}) \quad (4.7)$$

$$P_{++}^k(q, F) = 2^k q^{\frac{k}{6}} \left(\prod_{\beta=1}^k \cosh \frac{y_{\beta}}{2} \right) \prod_{n=1}^{\infty} \det(I + q^{2n} e^{\frac{iF}{2\pi}}) \quad (4.8)$$

$$P_{+-}^k(q, F) = 2^k q^{\frac{k}{6}} \left(\prod_{\beta=1}^k \sinh \frac{y_{\beta}}{2} \right) \prod_{n=1}^{\infty} \det(I - q^{2n} e^{\frac{iF}{2\pi}}) \quad (4.9)$$

where F is the Yang-Mills 2-form in the vector representation of $SO(2k)$ and I is the $2k \times 2k$ identity matrix. The factors of $\cosh(y_{\beta}/2)$ and $\sinh(y_{\beta}/2)^*$ in eqs. (4.8) and (4.9) come from the Chern characters of the Ramond vacuum: $P_{\pm\mp}^k(q, F)$ has a factor of $\text{Ch}(F, S^+) \pm \text{Ch}(F, S^-)$. Observe that eqs. (4.6)-(4.9) reduce to eqs. (4.1)-(4.4) when $y_{\beta} \equiv 0$.

Similarly, the generating function of the Chern characters of the representations of the transverse group generated from the world-sheet bosons is simply

$$P_B^N(q, R) = q^{-\frac{N}{6}} \prod_{n=1}^{\infty} \left[\det(I - q^{2n} e^{\frac{iR}{2\pi}}) \right]^{-1} \quad (4.10)$$

* Here and in the following, y_{β} ($\beta = 1, \dots, k$) are the skew eigenvalues of F , and x_{α} ($\alpha = 1, \dots, N$) are those of R .

where R is in the vector representation of $SO(2N)$ and I is the $2N \times 2N$ identity matrix.

Define now $T_1(q, F)$ to be the same combination of the $P_i^k(q, F)$'s as $T_1(q)$ is of the $P_i^k(q)$'s. The anomaly generating function is then

$$A(q, F, R) = \hat{A}(R) P_B(q, R) T_1(q, F) \quad (4.11)$$

Here 'F' stands generically for the complete gauge group [usually in this construction a product of $SO(2m)$ groups, but occasionally larger, such as the $E_8 \times E_8$ group of the heterotic string]. To derive the modular properties of $A(q, F, R)$, we first observe that the character valued partition functions take on a more familiar form when rewritten in terms of the skew eigenvalues of F and R . For example:

$$\begin{aligned} P_{--}^k(q, F) &= q^{-\frac{k}{12}} \prod_{n=1}^{\infty} \prod_{\beta=1}^k (1 + q^{2n-1} e^{y_{\beta}}) (1 + q^{2n-1} e^{-y_{\beta}}) \\ &= q^{-\frac{k}{12}} \prod_{n=1}^{\infty} \prod_{\beta=1}^k (1 + 2 \cosh y_{\beta} q^{2n-1} + q^{4n-2}) \\ &= \prod_{\beta=1}^k \frac{\mathcal{J}_3\left(\frac{y_{\beta}}{2\pi i} | \tau\right)}{\eta(\tau)} \end{aligned} \quad (4.12)$$

Similarly

$$P_{+-}^k(q, F) = \prod_{\beta=1}^k \frac{\mathcal{J}_4\left(\frac{y_{\beta}}{2\pi i} | \tau\right)}{\eta(\tau)} \quad (4.13)$$

$$P_{+0}^k(q, F) = \prod_{\beta=1}^k \frac{\mathcal{J}_2\left(\frac{y_{\beta}}{2\pi i} | \tau\right)}{\eta(\tau)} \quad (4.14)$$

$$P_{++}^k(q, F) = \prod_{\beta=1}^k \frac{\mathcal{J}_1\left(\frac{y_{\beta}}{2\pi i} | \tau\right)}{\eta(\tau)} \quad (4.15)$$

and

$$P_B^N(q, R) = \prod_{\alpha=1}^N \left[\frac{2 \sinh(\frac{x_\alpha}{2})}{\vartheta_1(\frac{x_\alpha}{2\pi i} | \tau)} \eta(\tau) \right] \quad (4.16)$$

and hence

$$\hat{A}(R) P_B^N(q, R) = \prod_{\alpha=1}^N \frac{x_\alpha \eta(\tau)}{\vartheta_1(\frac{x_\alpha}{2\pi i} | \tau)} \quad (4.17)$$

From Appendix B we find that the θ -functions transform according to

$$\vartheta_i\left(\frac{v}{c\tau+d} \mid \frac{a\tau+b}{c\tau+d}\right) = \epsilon \sqrt{c\tau+d} \exp\left(\frac{i\pi c v^2}{c\tau+d}\right) \vartheta_j(v | \tau) \quad (4.18)$$

where ϵ is a phase and j depends in some complicated way on i , a , b , c , and d . As explained in Appendix A, the transformation of $\eta(\tau)$ cancels the weight factor $\sqrt{c\tau+d}$, so that the ratios θ_i/η transform into each other with weight zero, apart from phases. Fortunately, none of the details are very relevant: we know that the phase as well as the relation between i and j does not depend on v . The only difference between the transformation for $v \neq 0$ and $v = 0$ is the exponential factor, which is the same for all θ -functions, and can therefore be factored out of the entire expression. The gravitational factor in the anomaly (4.17) transforms with a weight factor $(c\tau+d)^{-N}$ (provided we scale x_α) and a phase. Putting this all together we see that apart from the exponential factor, $A(q, F, R)$ has modular weight $-N$, as anticipated in Section 2. The extra exponential factors combine into an exponential of $\text{Tr } F^2 - \text{Tr } R^2$, where $\text{Tr } F^2$ is the sum of the traces of the vector representations of all $SO(2m)$ groups. We find, therefore, that under a general modular transformation

$$\begin{aligned}
& A\left(q\left(\frac{a\tau+b}{c\tau+d}\right), \frac{F}{c\tau+d}, \frac{R}{c\tau+d}\right) \\
&= \exp\left[\frac{ic}{32\pi^3(c\tau+d)} (\text{Tr } F^2 - \text{Tr } R^2)\right] \\
&\quad \times (c\tau+d)^{-N} A(q, F, R)
\end{aligned}
\tag{4.19}$$

The same transformation property holds for fermions with twisted boundary conditions. One simply replaces in all formulae the four θ -functions θ_i by $\theta\left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right](v|\tau)$. As one can see from formula (A.8), the modular transformation properties of the θ -function generalize to arbitrary α and β . Therefore all previous arguments remain the same.

As we will see in the next sections, this modular transformation property (4.19) is responsible for the anomaly cancellation. It will be shown in Section 7 that the fact that $A(q, F, R)$ is only a modular function if $\text{Tr } F^2 = \text{Tr } R^2$ is related to Quillen's anomaly [5].

5. THE HITCH HIKERS GUIDE TO MODULAR FUNCTIONS

This section will be a very brief review of the properties of modular functions that are relevant to our work.

Let H denote the upper half complex τ -plane [$\text{Im } (\tau) > 0$] and let $H^* = H \cup \{i\infty\}$ be H with the point at $i\infty$ appended. The group $SL(2, \mathbb{Z})$ can be viewed as the set of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c, d are integers satisfying $ad - bc = 1$. This group defines a set of maps from H to H via

$$\tau \rightarrow \frac{a\tau+b}{c\tau+d} \tag{5.1}$$

Observe that $-\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has the same action as $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and so we must factor this equivalence out. The modular group $\Gamma \equiv SL(2, \mathbb{Z})/\mathbb{Z}_2$ is the space of distinct maps of

the form (5.1) from H to H . The modular group can be generated from the maps $\tau \rightarrow \tau + 1$ and $\tau \rightarrow -1/\tau$.

A modular function of weight w is a meromorphic function $f: H^* \rightarrow \mathbb{C}$ such that

$$f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{-w} f(\tau) \quad (5.2)$$

for any $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$. The map f is called a modular form (or an entire modular function) if it is in fact holomorphic on H^* (i.e. including the point $\tau = i\infty$).

A fundamental set for the modular group action on H is a subset F of H that (i) maps onto the whole of H under the action of Γ and (ii) F must have no proper subset that maps onto the whole of H under Γ .

A set $\mathcal{F}$ is called a fundamental region if it contains a fundamental set and only points on the boundary of $\mathcal{F}$ are mapped into each other under the action of non-trivial elements of Γ [27].

The standard fundamental domain is shown in fig. 1. It is defined by

$$\mathcal{F} = \left\{ \tau : |\tau| \geq 1, -\frac{1}{2} < \text{Re } \tau < \frac{1}{2}, \text{Im } \tau > 0 \right\} \quad (5.3)$$

One should note that $\tau \rightarrow \tau + 1$ maps $\text{Re } (\tau) = -1/2$ onto $\text{Re } (\tau) = +1/2$, and that $\tau \rightarrow -1/\tau$ maps the arc from ρ to i onto the arc from ρ' to i . It is usual to add the point $\tau = i\infty$ to F , and we will adopt this convention^{*}.

It is sometimes convenient to define a set $D = \mathcal{F}/\Gamma = H^*/\Gamma$. This is the space obtained by identifying all points in $\mathcal{F}$ that are equivalent under Γ . Thus one glues $\text{Re } (\tau) = -1/2$ to $\text{Re } (\tau) = +1/2$, and the arc from ρ to i to the arc from ρ' to i (ρ and ρ' are thus viewed as the same point). Without the point $i\infty$ D has the topology

^{*}) This means that $\mathcal{F}$ is really the fundamental region for $H^* \cup Q$ where Q is the set of rational points on the $\text{Re } (\tau)$ axis.

of a disc, but since D includes $i\infty$ it has the topology of a 2-sphere. D is a 2-manifold except at the three cusp points ρ , i , and $i\infty$.

There is an obvious correspondence between modular functions on H^* and meromorphic functions $f: D \rightarrow \mathbb{C}$.

Define the index v_p of a (non-zero) modular function f at a point p by

$$v_p = \#(\text{zeroes of } f \text{ at } p) - \#(\text{poles of } f \text{ at } p) \quad (5.4)$$

(The zeros and poles are counted in their multiplicity.) A basic theorem in modular function theory relates the indices v_p and the weight w of a modular function f in the fundamental domain. Specifically

$$\left(\sum_{p \in \mathcal{F}_0} v_p \right) + v_\infty + \frac{1}{2} v_i + \frac{1}{3} v_\rho = \frac{w}{12} \quad (5.5)$$

where $\mathcal{F}_0$ is $\mathcal{F}$ with ρ , i and $i\infty$ excised, v_∞ is the index at $\tau = i\infty$, v_i and v_ρ are the indices at i and ρ (see fig. 1). The proof of eq. (5.5) is elementary, and may be found in any standard text book [27-29].

The idea is simple; consider

$$\oint \frac{f'}{f} dz \quad (5.6)$$

around the contour shown in fig. 2 but in the limit that $c, c' \rightarrow i\infty$ and the small arcs of the contour around ρ , i , and B go to zero. The points B and B' represent any pole or zero of f that lies on the boundary of $\mathcal{F}$. The residue theorem shows that (5.6) gives the first term in eq. (5.5). However, one can also evaluate (5.6) explicitly. The integration along c to c' gives $-v_\infty$ and the integrations along the small arcs around i , ρ , and ρ' give $-1/2 v_i - 1/3 v_\rho$. (Remember ρ and ρ' are identified.) The integration from c to ρ along $\text{Re}(\tau) = -1/2$ exactly cancels that along $\text{Re}(\tau) = +1/2$ from ρ' to c' simply because $f(\tau + 1) = f(\tau)$. The integration from ρ to i almost cancels that from i to ρ' for the same reason since $f(-1/\tau) =$

$\tau^w f(\tau)$. However, the weight factor is responsible for an uncanceled term, $-(w/\tau)$, in $\int f'/f dz$ from ρ to i . This yields the factor of $w/12$.

The first observation we can make from eq. (5.5) is that an entire modular function (i.e. no poles) must have even, non-negative weight. Furthermore, there are no entire non-vanishing modular functions of weight 2. Suppose that f is an entire modular function of weight zero. Let p be any point in $\mathcal{F}$. Then $g(\tau) = f(\tau) - f(p)$ is also an entire modular function of weight zero, but $g(\tau)$ has a zero at $\tau = p$. This contradicts eq. (5.5) unless $g(\tau) \equiv 0$, for all $\tau \in \mathcal{F}$. Thus the only entire modular functions of weight zero are constant functions.

One can readily construct examples of entire modular functions of weight greater than two on H^* . Consider

$$G_k(\tau) \equiv \sum'_{m,n \in \mathbb{Z}} (m\tau + n)^{-k} \quad (5.7)$$

where k is a positive integer and Σ' denotes the sum over all m and n , except $m = n = 0$. These functions are known as the Eisenstein series [27-29]. The sum in eq. (5.7) is only well defined when it is absolutely convergent, that is for $k \geq 3$, and it defines a holomorphic function on H [27-29]. For odd values of k , $G_k(\tau) \equiv 0$ since the sum over positive and negative values of m and n cancel. At $\tau = i\infty$

$$G_{2k}(i\infty) = 2 \zeta(2k) = -\frac{(2\pi i)^{2k}}{(2k)!} B_{2k} \quad (5.8)$$

and so $G_{2k}(\tau)$ is holomorphic on H^* .

By reordering the sum in eq. (5.7) it is clear that $G_{2k}(\tau + 1) = G_{2k}(\tau)$, for $k \geq 2$. Similarly, $G_{2k}(-1/\tau) = \tau^{2k} G_{2k}(\tau)$ and hence

$$G_{2k}\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^{2k} G_{2k}(\tau) \quad (5.9)$$

The Eisenstein function G_{2k} is thus an entire modular function of weight $2k$.

Since $G_{2k}(\tau + 1) = G_{2k}(\tau)$ it has a Fourier series. Indeed, one can show that for $k \geq 2$ [27-29]:^{*}

$$G_{2k}(\tau) = G_{2k}(q^2) = \frac{2(2\pi i)^{2k}}{(2k-1)!} \left[-\frac{B_{2k}}{4k} + \sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^{2n} \right] \quad (5.10)$$

where $q = e^{i\pi\tau}$, and

$$\sigma_k(n) = \sum_{d|n, d>0} d^k \quad (5.11)$$

i.e. $\sigma_k(n)$ is the sum of the k^{th} powers of the divisors of n .

For example

$$\begin{aligned} G_4(\tau) &= \frac{\pi^4}{45} \left[1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^{2n} \right] \\ G_6(\tau) &= \frac{2\pi^6}{945} \left[1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) q^{2n} \right] \\ G_8(\tau) &= \frac{16\pi^8}{4725} \left[1 + 480 \sum_{n=1}^{\infty} \sigma_7(n) q^{2n} \right] \end{aligned} \quad (5.12)$$

The functions G_4 and G_8 are, up to an overall normalization, the partition functions for the E_8 root lattice and the $E_8 \times E_8$ root lattice respectively [28].

Equation (5.5) shows that G_4 has exactly one zero, and this lies at $\tau = \rho$. Similarly G_6 has a single zero at $\tau = i$ and G_8 has a double zero at $\tau = \rho$. It follows that G_8/G_4^2 is an entire modular function of weight zero and hence a constant. The constant may be evaluated from eqs. (5.12) to yield

^{*} In refs. [1] and [2] we used a more convenient but less conventional definition of G_{2k} , in which the prefactor was absent.

$$G_8 = \frac{144}{21} G_4^2 \quad (5.13)$$

Define a function

$$\Delta = \frac{675}{256 \pi^{12}} [20 G_4^3 - 49 G_6^2] \quad (5.14)$$

From eqs. (5.12) we see that Δ vanishes at $q = 0$ or $\tau = i\infty$, and since Δ is an entire modular function of weight 12, $\tau = i\infty$ must be its only zero. Indeed one can show that Δ^{-1} is the partition function of the Veneziano model:

$$\Delta = q^2 \prod_{n=1}^{\infty} (1 - q^{2n})^{24} \quad (5.15)$$

Finally, define a function

$$j(\tau) = \frac{3^6 5^3}{\pi^{12}} \frac{G_4^3}{\Delta} \quad (5.16)$$

This function has weight zero, i.e. it is a modular invariant. It has a triple zero at q , and a simple pole at $\tau = i\infty$.

The first few terms in the Fourier series for $j(\tau)$ are

$$j(\tau) = j(q^2) = \frac{1}{q^2} + 744 + 196884 q^2 + 21493760 q^4 + \dots \quad (5.17)$$

The function $j - 744$ is the partition function of the Leech lattice.

Let $P(j)$ be any polynomial in $j(\tau)$. This is also a modular invariant, whose only poles lie at $\tau = i\infty$ or $q = 0$. Observe that under a modular transformation $\tau \rightarrow (a\tau + b)/(c\tau + d)$, $d/d\tau \rightarrow (c\tau + d)^2 (d/d\tau)$ and hence $(d/d\tau) [P(j)]$ is a modular function of weight 2. Let $f(\tau)$ be any modular function of weight 2, whose only

poles lie at $\tau = i\infty$ or $q = 0$. Then by suitable choice of the polynomial $P(j)$ one can arrange that $f(\tau) - (d/d\tau) P(j)$ is a modular function of weight 2 with no poles at all, and therefore zero. Thus we have shown that

$$f(\tau) = \frac{d}{d\tau} P(\tau) \quad (5.18)$$

From this we can conclude that if f is a modular function of weight 2 whose only poles lie at $q = 0$, then the coefficient of the q^0 term in the Fourier transform of $f(\tau)$ is identically zero. This is the basis of anomaly cancellation.

There is another more general argument by which we can arrive at this conclusion. Suppose $f(\tau)$ is a modular function of weight 2 whose only poles lie at some point p in $\mathcal{F}$. Choose a base point $\tau_0 \neq p$ in $\mathcal{F}$, and define

$$F(\tau) = \int_{\tau_0}^{\tau} f(\tau') d\tau' \quad (5.19)$$

where the integral is taken along any curve γ from τ_0 to τ that avoids p . Since $\mathcal{F}$ has the topology of S^2 , $\mathcal{F}/(p)$ has the topology of a disc, and is therefore simply connected. Thus the definition of $F(\tau)$ is independent of the path γ .

Under a modular transformation $s(\tau) = (a\tau + b)/(c\tau + d)$

$$F(s(\tau)) = \int_{\tau_0}^{s(\tau)} f(\tau') d\tau'$$

However, $f(\tau') d\tau'$ is modular invariant and is therefore equal to $f[s^{-1}(\tau')] d[s^{-1}(\tau')]$. Changing variables to $z = s^{-1}(\tau')$, we have

$$\begin{aligned} F(s(\tau)) &= \int_{s^{-1}(\tau_0)}^{\tau} f(z) dz \\ &= F(\tau) + \int_{s^{-1}(\tau_0)}^{\tau_0} f(z) dz \end{aligned}$$

The second integral vanishes since it is the integral of an analytic function over a closed curve in the simply connected domain $\mathcal{F}/\{p\}$. Thus we have explicitly constructed a function $F(\tau)$ such that $f(\tau) = (d/d\tau) F(\tau)$.

To conclude this section we discuss the anomalous Eisenstein function G_2 . Its anomaly is closely related to the Quillen anomaly. We remarked earlier that the series (5.7) is not absolutely convergent for $k = 1$, and is thus ill defined. However, we can still define $G_2(\tau)$ by specifying the order of summation:

$$G_2(\tau) = 2\zeta(2) + 2 \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} (m\tau+n)^{-2} \quad (5.20)$$

This is a holomorphic function on H^* [27]. However, since we have to reorder the sum in order to prove modular invariance, one expects that it is not modular invariant. [Indeed, we deduced from eq. (5.5) that it cannot be modular]. In order to define an 'almost' modular function of weight 2 related to $G_2(\tau)$, one can define [27]:

$$\Phi(\tau, s) = \sum'_{m, n \in \mathbb{Z}} (m\tau+n)^{-2} |m\tau+n|^{-s} \quad (5.21)$$

For $\text{Re}(s) > 0$ this is absolutely convergent and transforms with a modular weight $(c\tau+d)^2 |c\tau+d|^s$. One can analytically continue $\Phi(\tau, s)$ in s to $s = 0$ [27]. Define

$$\hat{G}_2(\tau) = \lim_{s \rightarrow 0} \Phi(\tau, s) \quad (5.22)$$

Thus one is ζ -function regularizing $G_2(\tau)$, and as one might expect, this preserves the modular invariance:

$$\hat{G}_2\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2 \hat{G}_2(\tau) \quad (5.23)$$

However, $\hat{G}_2(\tau)$ is not a holomorphic function of τ . Indeed [27]

$$\hat{G}_2(\tau) = G_2(\tau) - \frac{\pi}{\text{Im}(\tau)} \quad (5.24)$$

From this it follows that the holomorphic function $G_2(\tau)$ transforms under a modular transformation according to

$$G_2\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2 G_2(\tau) - 2\pi i c (c\tau+d) \quad (5.25)$$

Thus $G_2(\tau+1) = G_2(\tau)$, and so it has a Fourier series. One finds that [27, 29]

$$G_2(\tau) = G_2(q^2) = \frac{\pi^2}{3} \left[1 - 24 \sum_{n=1}^{\infty} \sigma_1(n) q^{2n} \right] \quad (5.26)$$

A final, rather amusing identity is obtained by defining

$$\begin{aligned} \psi(\tau) &= -\frac{1}{4\pi i} \left\{ \frac{\pi^2}{3} \left[\tau - 24 \sum_{n=1}^{\infty} \sigma_1(n) \frac{1}{2\pi i n} q^{2n} \right] \right\} \\ &= \frac{\pi i \tau}{12} - \sum_{n=1}^{\infty} \sigma_1(n) q^{2n} \end{aligned} \quad (5.27)$$

Observe that $(d/d\tau) \psi(\tau) = -4\pi i G_2(\tau)$. More surprisingly, we have [27, 29]:

$$\eta(\tau) = q^{\frac{1}{12}} \prod_{n=1}^{\infty} (1 - q^{2n}) = e^{\psi(\tau)} \quad (5.28)$$

The important thing to remember about $G_2(\tau)$ is that it is ill defined. Using eq. (5.20) it can be defined so that it is holomorphic, but it is not a modular function; there is an anomalous term in its modular transformation given by eq. (5.25). On the other hand, by using a form of ζ -function regularization we can obtain a function, $\hat{G}_2(\tau)$, of modular weight 2, but this function is not holomorphic. It is from this dichotomy that one finds the relation between the Quillen anomaly and the $\text{Tr}(F^2) - \text{Tr}(R^2)$ factor in the field-theory anomaly.

6. ANOMALY CANCELLATION

The anomaly is obtained by extracting the $(2N + 4)$ -form from the coefficient of q^0 in $A(q, F, R)$. Instead of doing this, consider your favourite $(2N + 4)$ -form in the anomaly. Let $f(\tau)$ be the coefficient of this $(2N + 4)$ -form in $A(q, F, R)$. Since the anomaly generating function is a sum of products of θ -functions, we see that $f(\tau)$ is holomorphic on H , with possible poles at $\tau = i\infty$ or $q = 0$.

Suppose now that $\text{Tr}(F^2) = \text{Tr}(R^2)$, then eq. (4.19) implies

$$A\left(q\left(\frac{a\tau+b}{c\tau+d}\right), \frac{F}{c\tau+d}, \frac{R}{c\tau+d}\right) = (c\tau+d)^{-N} A(q, F, R) \quad (6.1)$$

From this it follows that

$$f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{N+2} f(\tau) \quad (6.2)$$

where the factor of $(c\tau + d)^{N+2}$ comes from the fact that f is the coefficient of a $(2N + 4)$ -form.

From the previous section there is a polynomial $P(j)$ such that

$$f(\tau) = \frac{d}{d\tau} P(j) \quad (6.3)$$

and hence the q^0 term in $f(\tau)$ vanishes, that is the anomaly coefficient vanishes identically. Consequently the contribution of all the space-time fermions to the field-theory anomaly must have a factor $\text{Tr}(F^2) - \text{Tr}(R^2)$. Such an anomaly may be cancelled by the simplest Green-Schwarz mechanism.

Another, rather more suggestive way, of arriving at the same conclusion is to consider

$$\tilde{A}(q, F, R) = \exp\left[\frac{i}{64\pi^4} G_2(\tau) [\text{Tr} F^2 - \text{Tr} R^2]\right] A(q, F, R) \quad (6.4)$$

From (4.19) and (5.25) one finds that

$$\begin{aligned} \tilde{A}\left(q\left(\frac{a\tau+b}{c\tau+d}\right), \frac{F}{c\tau+d}, \frac{R}{c\tau+d}\right) \\ = (c\tau+d)^{-N} \tilde{A}(q, F, R) \end{aligned} \quad (6.5)$$

for any value of $\text{Tr}(F^2) - \text{Tr}(R^2)$. Therefore, by the foregoing argument, the coefficients of all the $(2N + 4)$ -forms in the q^0 term of $\tilde{A}(q, F, R)$ vanish identically. One then sees from eq. (6.4) that the anomaly term in $A(q, F, R)$ must have the factor of $\text{Tr}(F^2) - \text{Tr}(R^2)$.

The fact that $\tilde{A}(q, F, R)$ has no anomaly term leads us to speculate that it is the anomaly generating function for the field theory including the contribution of the antisymmetric tensor field $B_{\mu\nu}$.

7. THETA FUNCTIONS, DETERMINANTS, AND QUILLENS ANOMALY

In view of the results of the previous section, it is perhaps more surprising that the field-theory limit of a modular invariant string has an anomaly at all, rather than that the anomaly can be cancelled. If $A(q, F, R)$ had the modular transformation property (6.1) as opposed to eq. (4.19) then indeed the field-theory anomaly would vanish. It is only because of the anomalous exponential term in eq. (4.19) that the field-theory anomaly is present, and it is precisely because this anomalous exponential involves $\text{Tr}(F^2) - \text{Tr}(R^2)$ that the anomaly can be cancelled by the simplest Green-Schwarz mechanism. As was seen in Section 4, the anomalous exponential in the modular transformation of $A(q, F, R)$ is a direct result of the modular properties of θ -functions [see eq. (4.18)]. We therefore investigate the source of this anomalous exponential in the transformation of θ -functions.

The relationship between θ -functions and fermion determinants on the world sheet of the string has been discussed by a number of authors [18-20]. Here we will follow the discussion of ref. [20].

Consider a complex chiral spinor Ψ on the 2-torus. Parametrize this torus with two coordinates $0 \leq \sigma_1, \sigma_2 \leq 2\pi$, and let the metric be

$$ds^2 = |d\sigma_1 + \tau d\sigma_2|^2 \quad (7.1)$$

where τ is a complex Teichmüller parameter with $\text{Im}(\tau) > 0$. In this coordinate system, the chiral Dirac operator is

$$\not{D} = \gamma^0 \left[\frac{i}{\text{Im} \tau} \left(\frac{\partial}{\partial \sigma_2} + \tau \frac{\partial}{\partial \sigma_1} \right) \right] \quad (7.2)$$

Impose the following twisted boundary conditions on Ψ :

$$\begin{aligned} \psi(\sigma_1 + 2\pi, \sigma_2) &= -e^{2\pi i \alpha} \psi(\sigma_1, \sigma_2) \\ \psi(\sigma_1, \sigma_2 + 2\pi) &= -e^{2\pi i \beta} \psi(\sigma_1, \sigma_2) \end{aligned}$$

Then the chiral determinant, $\det(\alpha, \beta)$, of the Dirac operator on a complex fermion with these boundary conditions is given by [20]

$$\det(\alpha, \beta) = \frac{\vartheta \left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right] (0|\tau)}{\eta(\tau)} \quad (7.4)$$

(the θ -function is defined in Appendix A). This identity can be proved by use of the relationship between the path integral formalism and the canonical formalism [30]. We will essentially prove this result below. However, our present interest is to consider the modular behaviour of the θ -functions with two different values of β since this corresponds to varying the background fields [see eq. (A.2), and recall

that v was taken to be proportional to the skew eigenvalues of F or R]. We will relate the anomalous modular behaviour to an anomaly in $\det(\alpha, \beta)$.

From eq. (A.14)

$$\begin{aligned} \log \left[\frac{\mathcal{J} \left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right] (0|\tau)}{\mathcal{J} \left[\begin{smallmatrix} \alpha \\ \gamma \end{smallmatrix} \right] (0|\tau)} \right] &= 2\pi i \alpha (\beta - \gamma) + \sum_{n=1}^{\infty} \left\{ \log [1 + q^{2(n+\alpha)-1} e^{2\pi i \beta}] \right. \\ &\quad + \log [1 + q^{2(n-\alpha)-1} e^{-2\pi i \beta}] \\ &\quad - \log [1 + q^{2(n+\alpha)-1} e^{2\pi i \gamma}] \\ &\quad \left. - \log [1 + q^{2(n-\alpha)-1} e^{-2\pi i \gamma}] \right\} \end{aligned} \quad (7.5)$$

$$\begin{aligned} &= 2\pi i \alpha (\beta - \gamma) + \sum_{m=1}^{\infty} \frac{(-1)^m}{m} \frac{q^m}{1-q^{2m}} \left\{ q^{2\alpha m} (e^{2\pi i m \beta} - e^{2\pi i m \gamma}) \right. \\ &\quad \left. + q^{-2\alpha m} (e^{-2\pi i m \beta} - e^{-2\pi i m \gamma}) \right\} \end{aligned} \quad (7.6)$$

where we have expanded the logarithms in eq. (7.5), and performed the sum over n to obtain eq. (7.6). However, it is more convenient to keep the expanded form of eq. (7.6), and furthermore expand the exponentials, including $q^{2\alpha m}$, in the brackets.

The result is

$$\begin{aligned}
& 2\pi i \alpha (\beta - \gamma) + 2 \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{(-1)^m}{m} q^{(2n-1)m} (-1)^k \frac{1}{2k!} \\
& \quad \cdot \left\{ [2\pi m (\alpha\tau + \beta)]^{2k} - [2\pi m (\alpha\tau + \gamma)]^{2k} \right\} \\
& = 2\pi i \alpha (\beta - \gamma) + \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{2(2\pi i)^{2k}}{(2k-1)!} m^{2k-1} [(-q)^{mn} - (-q)^{2mn}] \\
& \quad \cdot \frac{1}{2k} [(\alpha\tau + \beta)^{2k} - (\alpha\tau + \gamma)^{2k}] \\
& = 2\pi i \alpha (\beta - \gamma) + \sum_{k=1}^{\infty} \frac{2(2\pi i)^{2k}}{(2k-1)!} \frac{1}{2k} [(\alpha\tau + \beta)^{2k} - (\alpha\tau + \gamma)^{2k}] \\
& \quad \cdot \left\{ \sum_{n=1}^{\infty} \sigma_{2k-1}(n) [(-q)^n - q^{2n}] \right\}
\end{aligned}$$

Hence

$$\begin{aligned}
& \log \left[\frac{\vartheta \left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix} \right] (0|\tau)}{\vartheta \left[\begin{smallmatrix} \alpha \\ \gamma \end{smallmatrix} \right] (0|\tau)} \right] \\
& = 2\pi i \alpha (\beta - \gamma) - \sum_{k=1}^{\infty} \frac{1}{2k} [(\alpha\tau + \beta)^{2k} - (\alpha\tau + \gamma)^{2k}] [\mathcal{G}_{2k}(q^2) - \mathcal{G}_{2k}(-q)] \quad (7.7)
\end{aligned}$$

where we have used eqs. (5.10) and (5.26). For the Jacobi θ -functions some simplifications can be made in the foregoing calculation to give

$$\begin{aligned}
\frac{\vartheta_1(v|\tau)}{v\vartheta_1'(0|\tau)} &= \exp \left\{ - \sum_{k=1}^{\infty} \frac{1}{2k} v^{2k} G_{2k}(q^2) \right\} \\
\frac{\vartheta_2(v|\tau)}{\vartheta_2(0|\tau)} &= \exp \left\{ \sum_{k=1}^{\infty} \frac{1}{2k} v^{2k} [G_{2k}(q^2) - 2^{2k} G_{2k}(q^4)] \right\} \\
\frac{\vartheta_3(v|\tau)}{\vartheta_3(0|\tau)} &= \exp \left\{ \sum_{k=1}^{\infty} \frac{1}{2k} v^{2k} [G_{2k}(q^2) - G_{2k}(-q)] \right\} \\
\frac{\vartheta_4(v|\tau)}{\vartheta_4(0|\tau)} &= \exp \left\{ \sum_{k=1}^{\infty} \frac{1}{2k} v^{2k} [G_{2k}(q^2) - G_{2k}(q)] \right\}
\end{aligned}$$

(7.8)

where we have used $v\vartheta_1'(0|\tau)$ in the first denominator since $\vartheta_1(0|\tau) \equiv 0$. Although eqs. (7.8) can be derived directly from eq. (7.7) it is more convenient to redo the calculation starting from formulae for $\log \vartheta$ given in ref. [31].

These equations enable one to readily understand the modular properties of $\vartheta_n(v|\tau)$. Observe that

$$\begin{aligned}
G_{2k}(q) &= G_{2k}\left(\frac{1}{2}\tau\right) \\
G_{2k}(-q) &= G_{2k}\left(\frac{1}{2}(\tau+1)\right) \\
G_{2k}(q^4) &= G_{2k}(2\tau)
\end{aligned}$$

Therefore, for $k \geq 2$, under $\tau \rightarrow \tau + 1$, $G_{2k}(q) \leftrightarrow G_{2k}(-q)$ and $G_{2k}(q^4) \rightarrow G_{2k}(q^4)$.

However, for $k \geq 2$,

$$\begin{aligned}
G_{2k}(q(-1/\tau)) &= G_{2k}\left(-\frac{1}{2\tau}\right) \\
&= (2\tau)^{2k} G_{2k}(2\tau) \\
&= 2^{2k} G_{2k}(q^4)
\end{aligned}$$

and

$$\begin{aligned}
 G_{2k}(-q(-1/\tau)) &= G_{2k}(\tau') \quad ; \quad \tau' = \frac{1}{\tau} \left(-\frac{1}{\tau} + 1 \right) \\
 &= (2\tau' - 1)^{-2k} G_{2k}(\tau'') \quad ; \quad \tau'' = \frac{\tau' - 1}{2\tau' - 1} = \frac{1}{2}(\tau + 1) \\
 &= \tau^{2k} G_{2k}\left(\frac{1}{2}(\tau + 1)\right) \\
 &= \tau^{2k} G_{2k}(-q)
 \end{aligned}$$

These results merely tell us how the θ -functions interchange under modular transformations; the weight factors appearing in the transformations of the Eisenstein functions are cancelled by rescaling $v \rightarrow v/(c\tau + d)$. The interesting transformation is that of G_2 . Using the same argument as above, we have

$$\begin{aligned}
 G_2(q^2(-1/\tau)) &= \tau^2 \left[G_2(q^2) - \frac{2\pi i}{\tau} \right] \\
 G_2(q(-1/\tau)) &= \tau^2 \left[4 G_2(q^4) - \frac{4\pi i}{\tau} \right] \\
 G_2(-q(-1/\tau)) &= \tau^2 \left[G_2(-q) - \frac{4\pi i}{\tau} \right] \\
 G_2(q^4(-1/\tau)) &= \tau^2 \left[\frac{1}{4} G_2(q) - \frac{\pi i}{\tau} \right]
 \end{aligned}$$

Using this in the modular transformation of eqs. (7.8) one finds that under $\tau \rightarrow -1/\tau$ all the ratios of θ -functions are multiplied by $\exp(i\pi v^2/\tau)$. Thus, the anomalous phase in the transformation of the θ -functions is precisely due to the lack of modular invariance of the holomorphic function $G_2(\tau)$.

Returning to eq. (7.7), we have

$$\frac{\vartheta\left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right](0|\tau)}{\vartheta\left[\begin{smallmatrix} \alpha \\ \gamma \end{smallmatrix}\right](0|\tau)} = e^{2\pi i \alpha(\beta-\gamma)} \exp \left\{ - \sum_{k=1}^{\infty} \sum_{m,n} \frac{1}{2k} [(\alpha\tau+\beta)^{2k} - (\alpha\tau+\gamma)^{2k}] \right. \\ \left. \cdot \left[\frac{1}{(m\tau+n)^{2k}} - \frac{1}{(\frac{m}{2}(\tau+1)+n)^{2k}} \right] \right\} \quad (7.9)$$

where $\Gamma'_{m,n}$ is defined according to eq. (5.20) for $k=1$. Summing over k one obtains

$$\frac{\vartheta\left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right](0|\tau)}{\vartheta\left[\begin{smallmatrix} \alpha \\ \gamma \end{smallmatrix}\right](0|\tau)} = e^{2\pi i \alpha(\beta-\gamma)} \prod_{m,n} \frac{[(m+\alpha+\frac{1}{2})\tau + (n+\beta+\frac{1}{2})]}{[(m+\alpha+\frac{1}{2})\tau + (n+\gamma+\frac{1}{2})]} \\ = e^{2\pi i \alpha(\beta-\gamma)} \frac{\det [\nabla(\alpha, \beta)]}{\det [\nabla(\alpha, \gamma)]} \quad (7.10)$$

where $\det [\nabla(\alpha, \beta)]$ is the determinant of $\nabla = (\partial/\partial\sigma_2) + \tau(\partial/\partial\sigma_1)$ over modes with boundary conditions (7.3). One should note that eq. (7.9) requires $G_2(\tau)$ to be defined by eqs. (5.20) or (5.26). Therefore, eq. (7.10) implicitly assumes that the determinants have been holomorphically regularized.

Putting it all together, we see that the θ -functions in the anomaly generating function can all be interpreted as determinants of fermion operators in suitable background fields. Expanding such a determinant as an exponential of the trace of the logarithm of the operator, and further expanding the logarithm of the operator is precisely the same as performing the Eisenstein expansion of the θ -function.

Schematically:

$$\vartheta \sim \det (\nabla^2 - v^2)^{1/2} \\ = [\det \nabla] \exp \left\{ \text{Tr} \frac{1}{2} \left[\log \left(1 - \frac{v^2}{\nabla^2} \right) \right] \right\} \\ = [\det \nabla] \exp \left\{ \sum_{k=1}^{\infty} \frac{1}{2k} v^{2k} \text{Tr} \left(\frac{1}{\nabla^{2k}} \right) \right\}$$

and $G_{2k} = \text{tr}'(\bar{v}^{-2k})$, where the trace is taken over modes (7.3) with $\alpha = \beta = 1/2$. However, the foregoing is simply the Feynman diagram expansion of the determinant in a background field. The $G_2(\tau)$ term corresponds to the logarithmically divergent graph shown in fig. 3. As has been discussed in ref. [20], another way of regulating this graph, and defining the determinant, is to introduce a fermion of opposite chirality and consider the well-defined determinant $\det[(\bar{v} - v) \bar{v}]$, and then take a holomorphic square root (if possible). This corresponds to regulating $G_2(\tau)$ using the ζ -function method described in Section 4. The result is the non-holomorphic, modular invariant map $\hat{G}_2(\tau)$. It is not possible to take the holomorphic square root of such a determinant precisely because the graph in fig. 3 cannot be regulated using $\bar{v}$ in such a way that the result is holomorphically factorizable. There is a factor of $\pi/\text{Im}(\tau)$ in the definition of $\hat{G}_2(\tau)$. This is Quillen's anomaly [5].

Thus we have an ambiguity. The natural object that must arise from the string is a sum of chiral fermion determinants. One can choose to regulate them holomorphically, in which case the Feynman graph in fig. 3 is regulated in the same way as $G_2(\tau)$. The result is our anomaly generating function $A(q, F, R)$ and a non-zero field theory anomaly, but one that has a factor of $\text{Tr}(F^2) - \text{Tr}(R^2)$. If one chooses the non-holomorphic, but modular invariant regularization it is not clear to us what the answer represents, since all our earlier anomaly cancellation arguments rested upon analyticity. What we have shown, however, is that the anomaly and its factorizability are both direct consequences of the fact that one cannot regularize $G_2(\tau)$ in such a way as to be both a modular function and holomorphic. It is in this sense that Quillen's anomaly is responsible for the factorizable field theory anomaly.

Finally, we observe that it is possible for the string theory to circumvent the problem of regulating $G_2(\tau)$. If our speculation at the end of Section 6 is correct, the contribution of $B_{\mu\nu}$ to the string amplitude might simply cancel out the $G_2(\tau)$ terms in the Eisenstein expansion. This is currently under investigation.

8. TYPE II STRINGS

By a type II string (as opposed to a heterotic string) we mean any closed string that has both left- and right-moving space-time fermions. Repeating the analysis of Section 2 we can write down the following expression for its one-loop path integral

$$\frac{\int \frac{d^2\tau}{(\text{Im } \tau)^2} (\text{Im } \tau)^{-N} \mathcal{P}_B^N(q) \mathcal{P}_B^N(\bar{q}) \sum_{i=1}^4 \sum_{j=1}^4 \mathcal{P}_i^N(\bar{q}) \mathcal{P}_j^N(q) T_{ij}(q, \bar{q}) \quad (8.1)$$

Here T_{ij} are the partition functions of any states other than the space-time bosons and fermions. The only chiral states in the theory are in the sectors $i = 1, j = 2, 3, 4$ and $i = 2, 3, 4, j = 1$. The $i = 1, j = 1$ sector has as its ground state $(\psi_L - \psi_R)^2$, which yields an anti-self-dual plus a self-dual tensor, and other bosons. We can thus immediately write down the partition function of the states contributing to the anomaly:

$$\begin{aligned} \mathcal{P}_A(q, \bar{q}) = & \mathcal{P}_B^N(q) \sum_{j=2}^4 T_{1j}(q, \bar{q}) \mathcal{P}_j^N(q) \\ & + \mathcal{P}_B^N(\bar{q}) \sum_{j=2}^4 T_{ji}(q, \bar{q}) \mathcal{P}_j^N(\bar{q}) \end{aligned} \quad (8.2)$$

To avoid massive chiral states, $\mathcal{P}_A(q, \bar{q})$ should not contain any terms with equal powers of q and $\bar{q}$. The obvious way to achieve this is to have T_{ij} depend only on q , and T_{ji} only on $\bar{q}$. Although the uniqueness of this answer is less clear than it was for the heterotic string, any other solutions would require that only the *closed* string is Lorentz invariant, not its two sectors separately. This is not what one expects in string theory, and we will not consider this possibility.

The two terms in eq. (8.2) should be separately modular invariant with weight factors $(c\tau+d)^{-N}$ and $(c\bar{\tau}+d)^{-N}$, respectively. This is exactly the same problem we have considered before, apart from the fact that $\mathcal{P}_j^N$ represents space-time symmetries

rather than gauge symmetries. If they would be gauge symmetries with field strength 2-form F , the Atiyah-Singer expression for the anomaly would be

$$\begin{aligned} & (\text{Tr } F^2 + \text{Tr } F_L^2 - \text{Tr } R^2) X_L^{2N-2} \\ & + (\text{Tr } F^2 + \text{Tr } F_R^2 - \text{Tr } R^2) X_R^{2N-2} \end{aligned}$$

where F_L and F_R are gauge fields that might come from T_{ij} and T_{ji} , and X_L^{2N-2} and X_R^{2N-2} are $(2N-2)$ -forms. We can now convert the $SO(2N)$ gauge symmetries to space-time symmetries simply by putting $F = R$ [15]. In the known superstrings F_L and F_R are not present, and the anomalies cancel then completely, as a consequence of modular invariance.

From previous arguments, we know that putting $F = R$ gives the right answer for tensor-spinors. A more subtle issue is whether it gives the right answer for spinor-spinors, or self-dual tensors. In the next section we show that the answer is positive, apart from possible problems with the relative sign of the bosonic and fermionic contributions.

An interesting question is whether one can actually construct chiral superstrings with non-vanishing F_L or F_R .

9. COVARIANT VERSUS LIGHT CONE ANOMALIES

Up to now we have assumed that for gravitational anomalies one may use $SO(d-2)$ transverse Lorentz group representations, whereas previously one has always used Euclidean $SO(d)$ matrices plus ghost contributions. An instructive example is the tensor product of an $SO(d-2)$ vector and left spinor, $S_+ \otimes V$ (where $+$ denotes the chirality). We would write for the anomaly of these states

$$\text{Anomaly } (S_+ \otimes V) = \hat{A}(R_T) \text{Ch}(R_T, V) \quad (9.1)$$

where "T" means transverse. In the covariant approach, one would say that the field content of this combination is a left-handed gravitino plus a right-handed spinor. The gravitino anomaly is given by refs. [7, 8]:

$$\hat{A}(R_C) [Ch(R_C, V) - 1] \quad (9.2)$$

with "C" for covariant, i.e. R_T has values in $SO(d-2)$ and R_C in $SO(d)$. The -1 in eq. (9.2) is the net contribution of the ghosts. Adding also the contribution of the physical right-handed spinor, we get

$$Anomaly(S_+ \otimes V) = \hat{A}(R_C) [Ch(R_C, V) - 2] \quad (9.3)$$

Instead of proving anomaly cancellation for $SO(d)$, one may consider only its $SO(d-2)$ subgroup. Absence of $SO(d-2)$ anomalies is certainly necessary for absence of $SO(d)$ anomalies.

In general, it is not sufficient: subgroups have fewer basic traces, and hence more trace identities than the group in which they are embedded. Therefore it is easier to cancel anomalies. In this case the problem is the index of order $d-2$, which $SO(d)$ has but $SO(d-2)$ does not. For $d > 6$, this index is irrelevant because the anomaly is of order $1/2 d + 1 < d - 2$. Furthermore, we never had to use the trace identity for $\text{Tr } R^{d-2}$ [see eq. (B.12)] in order to prove anomaly cancellation. It is sufficient to express all traces in terms of traces over the vector representation. The $\text{Tr } R^{d-2}$ terms cancel then directly. Therefore, it is legitimate to use the transverse Lorentz group even for $d \leq 6$.

Therefore we may decompose the $SO(d)$ vector to $SO(d-2)$. Because Chern characters decompose in the same way, we get

$$Ch(R_C, V) = Ch(R_T, V) + 2$$

From eq. (3.16) we see that $\hat{A}(R_C) = \hat{A}(R_T)$, so that eq. (9.3) is indeed equal to eq. (9.1).

The generalization to higher spin requires an understanding of all their ghosts. This is not a trivial issue in field theory, but is well understood in string theory. In the (ungauged) partition function, the ghost contributions reduce the "covariant" partition function to the light cone partition function, as follows

$$\mathcal{P}_B^{\frac{1}{2}d}(q) \mathcal{P}_{ghost}(q) = \mathcal{P}_B^{\frac{1}{2}(d-2)}(q) \quad (9.4)$$

In the presence of a gravitational background, the first factor is modified to become character valued [as in eq. (4.10)], but in $SO(d)$ rather than $SO(d-2)$. The world sheet ghosts cannot couple to the background fields, so that the second factor remains unchanged. Expanding the first factor as in Section 7, we get

$$\mathcal{P}_B^{\frac{1}{2}d}(q, R) = \mathcal{P}_B^{\frac{1}{2}d}(q) \exp \sum_{k=1}^{\infty} c_{2k} G_{2k}(q^2) \text{Tr} R_C^{2k} \quad (9.5)$$

where c_{2k} is an irrelevant constant. The ghost contribution reduces d to $d-2$ on the right-hand side of eq. (9.5), as it does in eq. (9.4), and in the exponent we can replace R_C by R_T , because no zeroth order traces appear. We have now obtained the light cone result.

This analysis does not yet include self-dual tensors, which appear in type II strings, but not in heterotic strings. In our method they appear when spinors of the gauge group are converted into space-time spinors. To simplify the discussion we will consider a type II string with identical left- and right-moving sectors, and no world-sheet fermions in addition to those that have space-time indices. The massless sector is

$$(B + S^+) \otimes (B + S^+) = 2 B \otimes S^+ + \phi^+ + \text{non-chiral bosons} \quad (9.6)$$

where B denotes bosonic excitations and ϕ^+ is a self-dual tensor. The self-dual tensor anomaly may either be obtained from the Hirzebruch signature polynomial (3.17), or equivalently from

$$\begin{aligned} \text{Index}(\phi^+) &= -\hat{A}(R_G) \text{Ch}(R_G, S^+) \\ &= -\hat{A}(R_T) [\text{Ch}(R_T, S^+) + \text{Ch}(R_T, S^-)] \end{aligned} \quad (9.7)$$

In the last step we used the reduction of the $SO(d)$ spinor to $SO(d-2)$. The - sign (relative to the anomaly of S^+) is due to the fact that ϕ^+ is a boson. In the construction of the previous section, the chiral part of the theory is obtained by writing S^+ in terms of the Ramond vacua of the $(++)$ and $(+-)$ spin structures

$$\begin{aligned} S_2 &= \frac{1}{2} (S^{++} + S^{+-}) \\ S_1 &= \frac{1}{2} (S^{++} - S^{+-}) \end{aligned}$$

The massless part of the partition function is then

$$(B + \epsilon S_2 \pm S_1)^2$$

The factor ϵ is determined by the modular invariance of the partition function, and determines the relative sign of space-time bosons and fermions to loop diagrams. The sign in front of S_1 is irrelevant, because it only determines the overall chirality. In our calculation of the anomaly of this theory, we would regard B and ϵS_2 as being coupled to gauge fields F, and replace F by R_T after proving that all gauge anomalies cancel. This method yields the following expression for the anomaly

$$\hat{A}(R_T) [2 \text{Ch}(R_T, B) + 2\epsilon \text{Ch}(R_T, S)] \quad (9.8)$$

(the factors of two appear because both terms are cross-products). This is precisely what one would derive for the anomaly of eq. (9.6) using eq. (9.7), provided that $\epsilon = -1$. If $\epsilon = +1$, the anomaly cancellation still holds, but it does not correspond to the field theory one obtains from the "string".

Instead, it corresponds to a field theory where either one assigns opposite (i.e. wrong) statistics to the self-dual tensor (or the fermions), or opposite chirality. The latter is meaningful in field theory, but of course not in string theory. This statistics problem arises, for example, in the proposed, but never constructed, 18-dimensional string [32-34].

10. ANOMALY-FREE FIELD THEORIES

An intriguing consequence of the generality of our results is that one can use modular invariance as a tool to construct very non-trivial anomaly-free field theories. Because the string theory is used here only as a tool, its consistency is not terribly important. If the string theory has tachyons, they are not part of its massless field theory limit; if the string theory does not have conformal invariance, we may not be able to get out of light cone gauge, but it may still be possible to construct a covariant field theory^{*)}; if the string theory has chiral states with wrong statistics, they can always be interpreted in anomaly graphs in field theory as states with correct statistics, and opposite chirality.

Here we will construct some of these anomaly-free field theories, without any attempt at generality. We will first give a complete construction of all one-loop modular invariant theories, obtainable with periodic and antiperiodic fermions^{**)}. If one has N left-moving and $\bar{N}$ right-moving fermions, then there are $4^{N+\bar{N}}$ possible

*) As a general rule, however, field-theory limits of non-conformal string theories turn out to have some problems as well: typically, they contain gravitinos and other higher spin states, without being supersymmetric.

**) A similar analysis, but with the additional requirement of a correct spin-statistics relation, was given in ref. [13]. However, we disagree with the transformation of the (++) spin structure used in that paper.

ways of assigning one of the four spin structures to each of them. To each of these assignments corresponds a path integral over a torus, and the one-loop amplitude is some linear combination of these $4^{N+\bar{N}}$ path integrals. Under modular transformations, this $4^{N+\bar{N}}$ -dimensional space is divided into invariant subspaces. They are characterized by choosing some partition of the $N + \bar{N}$ fermions into four (or fewer) groups, assigning spin structure 1 [= (++)] to one of these groups and each of the six permutations of spin structures 2, 3, and 4 to the others. Thus each subspace is (at most) six-dimensional.

All six permutations can be generated by a sequence of modular transformations $\tau \rightarrow \tau + 1$ and $\tau \rightarrow 1/\tau$, as follows

$$\tau \rightarrow \tau + 1 \rightarrow 1 - \frac{1}{\tau} \rightarrow \frac{\tau}{\tau + 1} \rightarrow \frac{1}{1 - \tau} \rightarrow -\frac{1}{\tau} \rightarrow \tau \quad (10.1)$$

In this way, one can generate the coefficient in front of each of the six terms from the coefficient of just one of them. Furthermore, there are consistency conditions, owing to the fact that one should get the original term back without a phase after six transformations, and also because each term should be invariant if $\tau \rightarrow \tau + 1$ is applied twice, since $\tau \rightarrow \tau + 2$ respects all boundary conditions. These conditions depend only on the number of left(right)-moving fermions with a boundary condition i (in one of the six terms), denoted by ℓ_i ($\bar{\ell}_i$). Furthermore, it depends only on the differences $\ell_i - \bar{\ell}_i \equiv m_i$. The conditions are easy to obtain from the modular transformation properties of the partition functions P_i (A.15) (we remind the reader that P_i represents two fermions). One finds

$$\begin{aligned} m_i &= 0 \pmod{4} \\ 2m_i + 2m_j - m_k - m_l &= 0 \pmod{24} \\ (\text{i.e. } m_i - m_j &= 0 \pmod{8}) \\ (i, j, k, l &\text{ is a permutation of } 1, 2, 3, 4) \end{aligned} \quad (10.2)$$

Furthermore, if two or more of the m_i 's ($i = 1, \dots, 4$) vanish, $m_i = 0 \pmod{8}$.

There are two simple ways of combining two modular invariant theories to get new ones. The first and most obvious one is linear combination. The second is convolution: one takes one term from each of the two modular invariant combinations, multiplies them, and then uses the sequence of transformations (10.1) to produce a new modular invariant theory. In general, six new theories are produced this way because there are six independent ways of forming the first pair (some of these six might be identical). The sum of all these six theories is the product of the two original ones.

The following combinations are clearly solutions to conditions (10.2)

$$a) \quad P_2^N(a) P_3^N(b) P_4^N(c) + \text{permutations of } (2, 3, 4)$$

$$b) \quad P_1^{-2}(a) [P_2^4(b) P_3^4(c) - P_3^4(b) P_2^4(c) \\ + P_3^4(b) P_4^4(c) - P_4^4(b) P_3^4(c) \\ + P_2^4(b) P_4^4(c) - P_4^4(b) P_2^4(c)]$$

Here $P_i^N(a)$ denotes a group of $2N$ fermions with spin structure i ; because $m_i = \ell_i - \bar{\ell}_i$ there is an obvious interpretation of negative N . The arguments a, b, c refer to fixed subsets of the $N + \bar{N}$ fermions. It is easy to see that by convolutions with these modular invariant combinations we can always modify the m_i 's in such a way that they are all $0 \bmod 8$, and their sum is $0 \bmod 24$. It is now convenient to enlarge the basis, by allowing also combinations in which two sets a and b have the same boundary conditions. Now we can complete squares, e.g.

$$\sum_{i \neq j} P_i^N(a) P_j^N(b) = \frac{1}{2} \left\{ \left[\sum_i P_i^N(a) \right] \left[\sum_j P_j^N(b) \right] - \sum_i P_i^{2N}(a, b) \right\}$$

where $i = 2, 3, 4$, and we have omitted some signs which are present if $N = 8 \bmod 16$. The sums on the right-hand side clearly form a complete basis. If we extend the

sums to include also the odd spin structure we get Neveu-Schwarz-Ramond (NSR) combinations with GSO projections. Their precise definition is

$$(NSR)_{\pm}^{8p}(a) = \frac{1}{2} \left[P_3^{4p}(a) + (-1)^p P_4^{4p}(a) + (-1)^p P_2^{4p}(a) \pm P_1^{4p}(a) \right]$$

The $\pm$ sign corresponds to the two possible chiralities in the fermion sector. The addition of the P_1 term is not required by one-loop modular invariance (it is required by two-loop modular invariance and factorization), but no generality is lost if we allow it to have both signs. The last possibility for constructing modular invariant theories is thus

- c) any product of NSR models with $8p$ fermions each and a multiple of 24 fermions combined (as usual these numbers are the difference between left and right movers).

The three known tachyon-free ten-dimensional heterotic theories can be obtained as follows. We denote the group of 8 right movers as a , and split the 32 left-moving fermions into two groups b and c of 16 fermions each. Then the three theories are given by (we omit chirality indices)

$$\begin{aligned} Spin(32)/Z_2 & \quad (NSR)^{-8}(a) (NSR)^{32}(b, c) \\ E_8 \times E_8 & \quad (NSR)^{-8}(a) (NSR)^{16}(b) (NSR)^{16}(c) \\ O(16) \times O(16) & \quad [E_8 \times E_8] - [Spin(32)/Z_2] \\ & \quad - (NSR)^8(a, b) (NSR)^{16}(c) - (NSR)^8(a, c) (NSR)^{16}(b) \\ & \quad + (NSR)^{24}(a, b, c) \end{aligned}$$

In the $O(16) \times O(16)$ theory, the last combination does not have any massless fermions, and the third and fourth combinations have chiral fermions in the

trivially anomaly-free combination $(128)_L + (128)_R'$ for each of the two $O(16)$ groups. It follows that the anomaly of the theory is just the difference of those of $E_8 \times E_8$ and $\text{Spin}(32)/Z_2$; the rôle of the third and fourth combination is only to flip simultaneously the chirality and statistics of some of the fermions in the first line, without affecting their contribution to the anomaly.

The 18- and 26-dimensional anomaly-free theories discussed in refs. [32-34] are obtained as follows:

$$\begin{aligned} (NSR)^{-16} (NSR)^{16} & \quad (d=18) \\ (NSR)^{-24} & \quad (d=26) \end{aligned}$$

where the right movers are interpreted as space-time fermions and the left movers as gauge fermions.

The obvious way to construct type II strings is to consider products or convolutions of left- and right-moving NSR models:

$$(NSR)_+^{-8\ell} (NSR)_\pm^{8\ell}$$

or

$$(NSR)_\pm^0$$

The first type is modular invariant only in 8ℓ dimensions, whereas the second type exists in any dimension; it is obtained by assigning the same boundary conditions to left and right movers (this type of string theory was first discussed in ref. [19], and does not have chiral fermions or self-dual tensors). The first type, with the same chirality for left and right movers, has a chiral field-theory limit, with the following content in light cone gauge:

$$2 S^+ \otimes (NS)_\ell + \phi^+$$

where S^+ is a Weyl spinor, ψ^+ a self-dual tensor of the same chirality (i.e. contained in $S^+ \otimes S^+$), and $(NS)_\ell$ is the representation content of the ℓ^{th} level of the Neveu-Schwarz model without GSO projections (with the ground state $|0\rangle$ labelled level 0, $b_{-1/2}^1 |0\rangle$ level 1, etc.). As discussed in Section 9, the anomaly can be calculated directly in light cone gauge, and is given by

$$\hat{A}(R) [2 \text{Ch}(R, (NS)_\ell) - \text{Ch}(R, S^-) - \text{Ch}(R, S^+)]$$

For odd ℓ , the $8\ell + 4$ form in this expression should, according to our theorem, cancel completely. For even ℓ one will discover that the sign in front of the self-dual tensor contribution has to be switched in order to get a cancellation. This is a consequence of the fact that such string theories do not have the correct spin-statistics relation for bosons and fermions.

The 24-dimensional type II theory is a spectacular example. In this case $\ell=3$, and we get

$$(NS)_3 = \begin{array}{|c|} \hline \square \\ \hline \end{array} + \square + \square \otimes \square$$

The 15 independent curvature 14-forms in the anomaly cancel completely among just six different spins. For higher ℓ , the cancellations become increasingly more spectacular and require bizarre number-theoretical identities involving Bernoulli numbers. Unfortunately, it is not obvious how to make sense of these higher spin field theories.

The combinations (a) and (b) defined above have also some interesting applications. The first one can give rise to a gauge group $SO(2N)^3$, for $6N$ fermions. Owing to the identity

$$\vartheta_2(0|\tau) \vartheta_3(0|\tau) \vartheta_4(0|\tau) = [2\eta(\tau)]^3$$

the partition function is actually a constant. When the θ -functions have a non-zero first argument this identity does not hold any more. The spectrum consists of a ground state $(S^+ + S^-, 1, 1) + \text{permutations}$, plus excitations which are higher $SO(2N)$ spinors. Because there is no odd spin structure, this combination by itself cannot produce a chiral theory. But it can be multiplied with any already modular invariant theory to produce an additional anomaly-free $SO(2N)^3$ gauge group.

Combination (b) finds its simplest application in six dimensions, with the four odd spin-structure fermions as right movers. The left movers produce an $SO(8) \times SO(8)$ gauge group, with massless Weyl fermions in the representation $(8_s + 8_c + 8_v, 1) - (1, 8_s + 8_c + 8_v)$. This is manifestly free of gravitational anomalies. Furthermore, the leading $\text{Tr } F^4$ terms cancel as well. This is most easily seen by observing that $8_s + 8_c + 8_v + 2(1)$ is the $SO(8)$ decomposition of the (26) of F_4 , a group which does not have a fourth index.

Clearly we have only scratched the surface with these examples. Using building blocks (a), (b), and (c) one can easily construct infinite numbers of anomaly-free chiral theories in any dimension.

11. CONCLUSIONS AND FINAL REMARKS

Our main result can be summarized as follows. With surprising generality, one-loop modular invariance, and one-loop modular invariance alone, guarantees Green-Schwarz factorization of anomalies in fermionic string theories, formulated in light cone gauge. It is not relevant for this conclusion (although certainly for the existence of the theory outside the light cone gauge) whether there are conformal anomalies. The $(\text{Tr } F^2 - \text{Tr } R^2)$ terms which do not cancel reveal the string origin of the field theory: they can be traced back to the two-dimensional anomaly, and in particular to Quillen's holomorphic anomaly. Although our rigorous proof depends on the way the theory is constructed, we expect it to be valid for any theory that is modular invariant, not just in its partition function but in all loop graphs with external lines, as explained in Section 2.

The basis of our approach is the relation of the anomaly to the Atiyah-Singer indices of the modes of the string. Just as in field theory, there are many different ways of thinking of anomalies in string theory. In ref. [3], we show how the expressions which we have constructed in this paper can be obtained from a generalization of Fujikawa's method to string field theory. The direct evaluation of the anomaly string loop diagram has also been completed recently [35].

Ironically, the anomaly cancellation which started the current wave of interest in string theory was for open strings with Chan-Paton gauge group, and is therefore not directly included in our results. But just as it led very quickly to the discovery of the heterotic string, it is tempting to speculate that some of the new, and far more spectacular, anomaly-free theories, which we now know to exist in higher dimensions, correspond to new string theories as well. Indeed, that possibility has already been considered [32]. Such theories will inevitably have higher spin massless fields, because the tachyon mass shifts downward with increasing dimension. According to conventional wisdom such theories do not make sense, not even as massless field theories. This is probably not unrelated to the conformal anomalies of the corresponding string theories. These facts, as well as our understanding of the once seemingly miraculous anomaly cancellations, argue against the existence of such new string theories.

Another tempting speculation concerns a possible relation between heterotic strings and the Fischer-Griess monster group F_1 [36]. This is based on the appearance of the function $j(q^2)$, whose Taylor coefficients are sums of dimensions of monster group representations [37], in the anomaly generating functions. Indeed, apart from the $(\text{Tr } F^2 - \text{Tr } R^2)$ terms, all gauge and Lorentz 12-forms in the anomaly of the $E_8 \times E_8$, $\text{spin } (32)/Z_2$, and $O(16) \times O(16)$ heterotic string are proportional to $(d/d\tau) j(q^2)$ (this is not the case for the superstring, which owes its anomaly cancellation to the trivial weight-two modular function). Perhaps the heterotic string is truly a theory of everything, containing not just all of physics, but also all known mathematics.

Acknowledgements

We would like to thank L. Alvarez-Gaumé, P. Ginsparg and K. Pilch for discussions. N.P.W. would like to thank the Aspen Centre for Physics for its hospitality while some of this work was carried out.

APPENDIX A

θ-FUNCTIONS

In this appendix we collect some formulae for θ-functions. Their general definition is (see, for example, ref. [20]).

$$\vartheta\left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right](r|\tau) = \sum_{n=-\infty}^{\infty} \exp\left[i\pi(n+\alpha)^2\tau + 2\pi i(n+\alpha)(r+\beta)\right] \quad (\text{A.1})$$

The first argument is actually redundant, because

$$\vartheta\left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right](r|\tau) = \vartheta\left[\begin{smallmatrix}\alpha \\ \beta+r\end{smallmatrix}\right](0|\tau) \quad (\text{A.2})$$

but it will be useful in some applications. It is not difficult to derive the basic modular transformation properties of this function. For the transformation $\tau \rightarrow \tau + 1$ one derives in a straightforward way

$$\vartheta\left[\begin{smallmatrix}\alpha \\ \beta-\alpha+\frac{1}{2}\end{smallmatrix}\right](r|\tau+1) = e^{-i\pi(\alpha^2-\alpha)} \vartheta\left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right](r|\tau) \quad (\text{A.3})$$

To obtain the $\tau \rightarrow -1/\tau$ transformation rule, one observes that $\vartheta\left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right]$ is periodic in α , so that it can be expanded in a Fourier series

$$\vartheta\left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right](0|\tau) = \sum_{m=-\infty}^{\infty} e^{2\pi i m \alpha} a_m \quad (\text{A.4})$$

where

$$a_m = \sum_{n=-\infty}^{\infty} \int_0^1 d\alpha e^{-2\pi i m \alpha} e^{[i\pi(n+\alpha)^2\tau + 2\pi i(n+\alpha)\beta]} \quad (\text{A.5})$$

In the Fourier phase we may replace α by $\alpha + n$; then the sum and the integral can be combined to give an integral from $-\infty$ to $+\infty$. For $\text{Im } \tau > 0$ this is a well-defined Gaussian; evaluating it and substituting the result in eq. (A.4) we obtain

$$\mathcal{G}\left[\begin{smallmatrix} \beta \\ -\alpha \end{smallmatrix}\right](0|\frac{1}{\tau}) = \sqrt{-i\tau} e^{-2\pi i \beta \alpha} \mathcal{G}\left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right](0|\tau) \quad (\text{A.6})$$

where the square root must be chosen to lie in the right half plane. The transformation properties under arbitrary modular transformations

$$\tau \rightarrow \tilde{\tau} = \frac{a\tau + b}{c\tau + d}$$

can now be obtained by repeated use of eqs. (A.3) and (A.6). The result can be summarized by the following formula [20]:

$$\mathcal{G}\left[\begin{smallmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{smallmatrix}\right](0|\tilde{\tau}) = \varepsilon(\Lambda) e^{-i\pi\phi(\alpha, \beta, \Lambda)} \sqrt{c\tau + d} \mathcal{G}\left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right](0|\tau) \quad (\text{A.7})$$

where $\Lambda = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$\begin{bmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \frac{1}{2} \begin{bmatrix} cd \\ ab \end{bmatrix}$$

$$\phi(\alpha, \beta, \Lambda) = \alpha^2 db + \beta^2 ac - 2\alpha\beta bc - (\alpha d - \beta c)ab$$

The phase ε depends in a rather complicated way on Λ . It is usually simpler to use directly eqs. (A.3) and (A.6) to determine it. These transformation rules are for $v = 0$. Using eq. (A2) we can restore v , and derive with some straightforward algebra

$$\mathcal{G}\left[\begin{smallmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{smallmatrix}\right]\left(\frac{v}{c\tau+d} \middle| \frac{a\tau+b}{c\tau+d}\right) = \varepsilon(\Lambda) e^{i\pi\phi(\alpha, \beta, \Lambda)} e^{\frac{i\pi v^2 c}{c\tau+d}} \sqrt{c\tau+d} \mathcal{G}\left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right](v|\tau) \quad (\text{A.8})$$

It is important that neither the transformation of the arguments α , β , nor the first two phases acquire a v -dependence, and furthermore that the extra phase $\exp(i\pi c v^2/c\tau + d)$ does not depend on α and β .

The Jacobi θ -functions can now be defined as follows [38]:

$$\begin{aligned} \vartheta_1 &= \vartheta \left[\begin{smallmatrix} i \\ i \end{smallmatrix} \right] & \vartheta_2 &= \vartheta \left[\begin{smallmatrix} i \\ 0 \end{smallmatrix} \right] \\ \vartheta_3 &= \vartheta \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right] & \vartheta_4 &= \vartheta \left[\begin{smallmatrix} 0 \\ i \end{smallmatrix} \right] \end{aligned} \quad (\text{A.9})$$

Here we have suppressed the arguments $(v|\tau)$, which are identical on both sides of the equation. The transformation properties of these functions can be read off from eqs. (A.3), (A.6), and (A.8).

Another important function is the Dedekind η -function

$$\eta(\tau) = q^{\frac{i}{24}} \prod_{n=1}^{\infty} (1 - q^{2n}) \quad (\text{A.10})$$

(with, as usual, $q = e^{i\pi\tau}$). Its transformation properties follow from those of θ_1 because of the relation

$$\vartheta_1'(0|\tau) = 2[\eta(\tau)]^3 \quad (\text{A.11})$$

where $\vartheta_1'(0|\tau) \equiv (\partial/\partial v) \vartheta_1(v|\tau)|_{v=0}$. The basic transformations are

$$\begin{aligned} \eta(\tau+1) &= e^{\frac{i\pi}{24}} \eta(\tau) \\ \eta\left(-\frac{i}{\tau}\right) &= \sqrt{-i\tau} \eta(\tau) \end{aligned} \quad (\text{A.12})$$

The weight factors $\sqrt{-i\tau}$ or $\sqrt{i\tau + d}$ cancel in the transformations of the partition functions $P \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ (for two fermions).

$$\rho \begin{bmatrix} \alpha \\ \beta \end{bmatrix}(\tau|\tau) = \frac{\vartheta \begin{bmatrix} \alpha \\ \beta \end{bmatrix}(\tau|\tau)}{\eta(\tau)} \quad (\text{A.13})$$

We also define P_i ($i = 1, \dots, 4$), analogous to eqs. (A.9). The partition functions can be expressed in terms of an infinite product:

$$\rho \begin{bmatrix} \alpha \\ \beta \end{bmatrix}(0|\tau) = e^{2\pi i \alpha \beta} q^{\alpha^2 - \frac{1}{2}} \prod_{n=1}^{\infty} (1 + q^{2n-1+2\alpha} e^{2\pi i \beta}) (1 + q^{2n-1-2\alpha} e^{-2\pi i \beta}) \quad (\text{A.14})$$

with $0 \leq \alpha, \beta \leq 1$. Finally, we list the transformation properties of the P_i 's. If $\tau \rightarrow -1/\tau$

$$\begin{aligned} P_1 &\rightarrow e^{-\frac{i\pi}{2}} P_1 & P_3 &\rightarrow P_3 \\ P_2 &\rightarrow P_4 & P_4 &\rightarrow P_2 \end{aligned}$$

If $\tau \rightarrow \tau + 1$

$$\begin{aligned} P_1 &\rightarrow e^{\frac{i\pi}{6}} P_1 & P_3 &\rightarrow e^{-\frac{i\pi}{12}} P_4 \\ P_2 &\rightarrow e^{\frac{i\pi}{6}} P_2 & P_4 &\rightarrow e^{-\frac{i\pi}{12}} P_3 \end{aligned} \quad (\text{A.15})$$

in an obvious shorthand notation for formula (A.8). The v^2 -dependent phase has been omitted; the weight factor is not present.

The partition function for 2N bosons is defined as

$$P_B^N(q^2) = [\eta(q^2)]^{-2N} \quad (\text{A.16})$$

APPENDIX B

TRACE IDENTITIES

Here we will show how to derive identities for higher order traces in $SU(N)$, $SO(N)$, and $Sp(N)$. We begin by expanding eq. (3.9) directly in terms of $\text{Tr } F^n$ for the group $SU(N)$

$$\prod_{\ell=1}^{\infty} \exp \left[-\frac{(-x)^\ell}{\ell} \text{Ch}(\ell F) \right] = (1+x)^N \exp \left\{ \sum_{n=2}^{\infty} \frac{1}{n!} \frac{P_n^d(x)}{(1+x)^d} \text{Tr} \left(\frac{iF}{2\pi} \right)^n \right\} \quad (\text{B.1})$$

The index polynomials $P_n^d(x)$ are defined as follows:

$$\begin{aligned} P_n^d(x) &= -(1+x)^d \sum_{\ell=1}^{\infty} (-x)^\ell \ell^{n-1} \\ &= (1+x)^d \left(x \frac{d}{dx} \right)^{n-1} \frac{x}{1+x} \end{aligned} \quad (\text{B.2})$$

The factor $(1+x)^d$ is introduced for convenience: if a term of a given order in F is to be extracted from eq. (B.1), one can adjust the value of d in each factor in such a way that the prefactor $(1+x)^N$ is precisely cancelled. For $d \geq n$ these polynomials are of finite degree. The coefficient of x^k is then precisely the n^{th} index of the representation $[k]$ of $SU(d)$ (with a normalization so that $I_n([1]) = 1$):

$$P_n^d(x) = \sum_{k=0}^d I_n^d([k]) x^k \quad (\text{B.3})$$

A useful expression for the coefficients is

$$I_n^d([k]) = \sum_{\ell=1}^k (-1)^\ell \ell^{n-1} \dim([k-\ell])_{SU(d)} \quad (\text{B.4})$$

By construction, the product of several index polynomials depends only on the sum of their upper indices, not on these indices separately. (Of course the product

does depend on the lower indices separately.) The polynomials vanish at $x = 0$.

Furthermore, one can prove

$$P_n^d(-1) = -(d-1)! \delta_{nd} \quad (d \geq n) \quad (\text{B.5})$$

$$P_{2n}^d(1) = 2^{d+1} \frac{2^{2n}-1}{4n} B_{2n} \quad (\text{B.6})$$

$$P_{2n+1}^d(1) = 0 \quad (\text{B.7})$$

where B_{2n} is a Bernoulli number. The first of these relations follows directly from the definition (B.2); the last one from the fact that odd traces have opposite signs for complex conjugate representations. The second relation will be derived later, using spinor representations of $SO(N)$.

SU(N) trace identities

The SU(N) trace identities can now be derived from the fact that antisymmetric tensors of rank higher than N do not exist, so that eq. (B.1) should terminate, i.e. should be a finite polynomial in x . Consider the terms of order M in the expansion of eq. (B.1). They are given by

$$(1+x)^{N-M} \sum_{\{n_i, m_i\}_M} \prod_i \frac{1}{m_i!} \left\{ P_{n_i}^{m_i}(x) \frac{1}{n_i!} \text{Tr} \left(\frac{iF}{2\pi} \right)^{n_i} \right\}^{m_i} \quad (\text{B.8})$$

where the sum is over all partitions (n_i, m_i) with $\sum_i n_i m_i = M$. The maximum number of factors $(1+x)^{-1}$ has been absorbed in $P_{n_i}^{m_i}$ (i.e. $P_{n_i}^{m_i}$ is not a finite polynomial). Because the expression must be a finite polynomial, we know that the residue of the pole at $x = -1$ must vanish. Using eq. (B.5) we find then

$$\sum_{\{n_i, m_i\}_M} \prod_i \frac{1}{m_i!} \left(- \frac{\text{Tr} F}{n_i} \right)^{n_i m_i} = 0 \quad (M > N) \quad (\text{B.9})$$

[For example, in SU(3) we find $\text{Tr } F^4 = 1/2 (\text{Tr } F^2)^2$, $\text{Tr } F^5 = 5/6 (\text{Tr } F^3 \text{Tr } F^2)$ etc.]

SO(2M) trace identities

Identity (3.9) is of course valid for the vector representation of SO(2M), with all odd traces equal to zero [the same is true for SO(2M + 1) and Sp(2M)]. In SO(2M), the existence of the chiral trace $\{\tilde{M}\}$ [see eq. (3.14)] allows us to express $\text{Tr } F_{\nu}^{2M}$ in terms of $\{\tilde{M}\}^2$. This can be achieved by considering products of spinors. First of all, one can use the relations

$$S \otimes S = \frac{1}{2} \sum_{k=0}^N [k] \quad (SO(N), N \text{ odd})$$

and

(B.10)

$$(S^- + S^+) \otimes (S^- + S^+) = \sum_{k=0}^N [k] \quad (SO(N), N \text{ even})$$

to obtain the value of $p_{2n}^d(x)$ at $x = 1$ [see eq. (B.6)]. Then one can use

$$S^- \otimes S^+ = \frac{1}{4} \left\{ \sum_{e=0}^N (1 + (-1)^e) x^e [e] \right\}_{x=1} \quad (N = 4k+2)$$

and

(B.11)

$$S^- \otimes S^+ = \frac{1}{4} \left\{ \sum_{e=0}^N (1 - (-1)^e) x^e [e] \right\}_{x=1} \quad (N = 4k)$$

to derive the SO(N) trace relation

$$\sum_{\{n_i, m_i\}_N} \prod_i \frac{1}{n_i!} \left[-\frac{1}{n_i} \text{Tr} \left(\frac{iF}{2\pi} \right)^{n_i} \right]^{m_i} = \mp \left\{ \frac{\tilde{N}}{2} \right\}^2 \quad (B.12)$$

where the upper sign is for $N = 4k + 2$, and the lower one for $N = 4k$. The notation is as in eq. (B.9), i.e. $\{n, m\}_N$ denotes all partitions with $\sum_i m_i = N$. For example, in $SO(4)$ one gets

$$-\frac{1}{4} \text{Tr } F^4 + \frac{1}{8} (\text{Tr } F^2)^2 = (\text{Tr } F_{S^\pm})^2$$

REFERENCES

- [1] A.N. Schellekens and N.P. Warner, Anomalies and modular invariance in string theory, preprint CERN-TH 4464/86, to appear in Phys. Lett. B.
- [2] A.N. Schellekens and N.P. Warner, Anomaly cancellation and self dual lattices, preprint CERN-TH 4465/86, to appear in Phys. Lett. B.
- [3] K. Pilch, A.N. Schellekens and N.P. Warner, Path integral calculation of string anomalies, MIT preprint.
- [4] M.B. Green and J. Schwarz, Phys. Lett. 149B (1984) 117.
- [5] D. Quillen, Funk. Anal. i Prilozhen. 19 (1985) 37.
- [6] M.F. Atiyah and I.M. Singer, Ann. Math. 87 (1968) 485, 546, and 93 (1971) 119, 139.

M.F. Atiyah and G.B. Segal, Ann. Math. 87 (1968) 531.
- [7] L. Alvarez-Gaumé and E. Witten, Nucl. Phys. B234 (1983) 269.
- [8] L. Alvarez-Gaumé and P. Ginsparg, Ann. Phys. (NY) 161 (1985) 423.
- [9] O. Alvarez, I. Singer and B. Zumino, Commun. Math. Phys. 96 (1984) 409.
- [10] D. Gross, J. Harvey, E. Martinec and R. Rohm, Nucl. Phys. B256 (1985) 253; B267 (1986) 75.
- [11] L. Alvarez-Gaumé, P. Ginsparg, G. Moore and C. Vafa, An $O(16) \times O(16)$ heterotic string, Harvard preprint HUTP-86/A013 (1986).

L. Dixon and J. Harvey, String theories in ten dimensions without space-time supersymmetry, Princeton preprint (1986).
- [12] L. Dixon, J. Harvey, C. Vafa and E. Witten, Nucl. Phys. B261 (1985) 651; Strings on orbifolds II, Princeton preprint (1986).
- [13] H. Kawai, D.C. Lewellen and S.-H. H Tye, Classification of closed fermionic string models, Cornell preprint CLNS 86/732 (1986); Construction of four dimensional fermionic string models, CLNS 86/734 (1986).
- [14] K.S. Narain, Phys. Lett. 169B, (1986) 41.
- [15] A.N. Schellekens, Phys. Lett. B175 (1986) 41.
- [16] A.M. Polyakov, Phys. Lett. 103B (1981) 207, 211.

- [17] J. Polchinsky, Commun. Math. Phys. 104 (1986) 37.
- [18] E. d'Hoker and D.H. Phong, Loop amplitudes for the fermionic string, Columbia preprint CU-TP-323.
- [19] N. Seiberg and E. Witten, Spin structures in string theory, Princeton preprint (1986).
- [20] L. Alvarez-Gaumé, G. Moore and C. Vafa, Theta functions, modular invariance, and strings, Harvard preprint HUTP-86/A017 (1986).
- [21] J. Shapiro, Phys. Rev. D5 (1972) 1945.
- [22] J.H. Schwarz, Phys. Rep. 89 (1982) 223.
- [23] H. Freudenthal and H. de Vries, Linear Lie groups (Academic Press, New York, 1969).
- [24] S. Okubo and J. Patera, J. Math. Phys. 24 (1983) 2722 and 25 (1984) 219.
- [25] P. Frampton and T. Kephart, Phys. Rev. D28 (1983) 1010.
- [26] T. Eguchi, P. Gilkey and A. Hanson, Phys. Rep. 66 (1980) 213.
- [27] B. Schoeneberg, Elliptic modular functions (Springer, Berlin-Heidelberg, 1970).
- [28] J.-P. Serre, A course in arithmetic (Springer, Berlin-Heidelberg, 1973).
- [29] B. Rankin, Modular forms and functions (Cambridge University Press, Cambridge, 1977);
S. Lang, Elliptic functions (Addison-Wesley 1973).
- [30] L. Faddeev and A. Slavnov, Gauge fields: the quantum theory (Benjamin, New York, 1982).
- [31] M. Abramowitz and I. Stegun, Handbook of mathematical functions (Dover, New York, 1965).
- [32] J. Thierry-Mieg, Phys. Lett. 171B (1986) 163.
- [33] W. Nahm, Commun. Math. Phys. 105 (1986) 1.
- [34] F. Englert, H. Nicolai and A.N. Schellekens, Nucl. Phys. B274 (1986) 315.

- [35] A. Sugamoto and H. Suzuki, Gauge and gravitational anomalies in heterotic superstrings, KEK preprint 85-89 (1985); Role of modular invariance in evaluation of gauge and gravitational anomalies in heterotic string, KEK preprint 86-35 (1986).
- [36] R.L. Griess, Inv. Math. 69 (1982) 1.
- [37] J.G. Thompson, Bull. London Math. Soc. 11 (1979) 352.
J.H. Conway and S.P. Norton, Monstrous moonshine, Bull. London Math. Soc. 11 (1979) 308.
- [38] A. Erdelyi et al., Higher transcendental functions (McGraw Hill, New York, 1953).

Figure captions

Fig. 1 : The standard fundamental region for Γ in H .

Fig. 2 : Integration contour for (5.6).

Fig. 3 : The two-dimensional logarithmically divergent fermion loop graph responsible for modular anomalies or Quillen's anomaly.

Fig. 1

Fig. 2

Fig. 3